DOCUMENT RESUME

ED 415 621 EC 306 101

AUTHOR	Ysseldyke, Jim; Thurlow, Martha; Erickson, Ron; Gabrys, Robert; Haigh, John; Trimble, Scott; Gong, Brian
TITLE	A Comparison of State Assessment Systems in Kentucky and Maryland with a Focus on the Participation of Students with Disabilities. State Assessment Series: Maryland/Kentucky Report 1.
INSTITUTION	Maryland State Dept. of Education, Baltimore.; National Center on Educational Outcomes, Minneapolis, MN.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE	1996-11-00
NOTE	96p.
CONTRACT	H023F50008
AVAILABLE FROM	National Center on Educational Outcomes, University of Minnesota, 350 Elliott Hall, 75 East River Road, Minneapolis, MN 55455; telephone: 612-624-8561; fax: 612-624-0879; World Wide Web: http://www.coled. umn.edu/NCEO (document may be copied without charge, additional print copies, $10).
PUB TYPE	Reports - Descriptive (141) -- Reports - Research (143) -- Tests/Questionnaires (160)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Academic Standards; *Accountability; *Disabilities; *Educational Assessment; Educational Testing; Elementary Secondary Education; Evaluation Methods; Models; Outcomes of Education; Program Costs; State Programs; *Student Evaluation; *Student Participation; Test Use
IDENTIFIERS	Kentucky; Maryland; Testing Accommodations (Disabilities)

ABSTRACT

This report describes and compares the inclusive assessment and accountability systems of Maryland and Kentucky, with a focus on components, scoring methods, reporting practices, data use, participation of students with disabilities, testing accommodations, and implementation considerations (including cost and training issues). Maryland and Kentucky practices are then contrasted to those used in other states. Results of the study indicate that both states have multiple forms of assessment and explicit procedures for deciding who participates in the various assessments. In both Maryland and Kentucky, schools are held accountable for improved student performance. In Maryland, schools are expected to show progress toward state-defined long-term goals. In Kentucky, schools are expected to improve over baseline performance. Both states and most others permit accommodations in assessments. However, Maryland and Kentucky are among the very few states that have an alternate assessment system that permits participation by students with severe disabilities. The report concludes that the assessment and accountability systems used in Kentucky and Maryland serve as good models for other states. Appendices include samples of released items and scoring guides from the Maryland School Performance Assessment Program and from the Kentucky Instructional Results Information System. (Contains 12 references.) (CR)

ED 415 621

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

EC 306401

2

STATE ASSESSMENT SERIES
Maryland/Kentucky
Report 1

A Comparison of State Assessment Systems in Kentucky and Maryland

with a focus on the participation of students with disabilities

The Maryland-Kentucky-NCEO Assessment Project

Prepared By:

Jim Ysseldyke • Martha Thurlow • Ron Erickson
National Center on Educational Outcomes, University of Minnesota

Robert Gabrys • John Haigh
Maryland Department of Education

Scott Trimble • Brian Gong
Kentucky Department of Education

November 1996

The Maryland-Kentucky-NCEO Assessment Project encompasses a comprehensive array of research studies on assessment accommodations and alternate assessment approaches that facilitate inclusion of all students in statewide assessment programs. The planned research program include:

- comparative studies involving a qualitative examination of the assessment systems in Kentucky and Maryland;
- studies that involve secondary analyses of existing data bases to address critical technical and implementation issues in the assessment of students with disabilities; and
- experimental field studies that involve the collection of new data to address several technical issues crucial to the development and modification of state assessment policies.

This project is supported, in part, by a grant to the Maryland State Department of Education from the U.S. Department of Education, Office of Special Education Programs (Grant # H023F50008). Opinions expressed herein do not necessarily reflect those of the U.S. Department of Education or Offices within it.

Permission is granted to copy this document without charge. Additional print copies may be purchased for $10.00 from:

NATIONAL CENTER ON EDUCATIONAL OUTCOMES
University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612.624.8561 • Fax 612.624.0879
http://www.coled.umn.edu/NCEO

This document is available in alternative formats upon request.

Abstract

States currently are struggling with the development of assessment and accountability systems that are for all students. Two states, Maryland and Kentucky, have made major advances toward fully inclusive assessment systems. We describe and compare the systems used in the two states with a focus on components, scoring methods, reporting practices, data use, participation of students with disabilities, testing accommodations, and implementation considerations. Maryland and Kentucky practices are then contrasted to those used in the other states.

Nearly every state department of education is engaged in the specification of standards and the development of systems of assessment to use in making judgments about the extent to which students and schools are meeting high standards (AFT, 1996; Bond & Roeber, 1995). Major reform legislation (Goals 2000, School-to-Work, Improving America's Schools Act, and the Individuals with Disabilities Education Act) includes wording indicating that high standards and accompanying assessments are for all students. Two states, Maryland and Kentucky, have made major advances in the development of fully inclusive assessment and accountability systems. Maryland has about 99% of its students participating in the assessments, and Kentucky has 100%. In this paper, we describe and compare the assessment systems and practices in the two states. The purpose in documenting what is happening in these two states is both to clarify the systems and to provide models for other states.

Assessment Practices in Maryland

Overview

Maryland's School Performance Program (MSPP) was established in 1989 by the Maryland State Board of Education as the vehicle to move toward a high quality educational system for all of Maryland's students in the 21st century. Maryland's focus on school performance and standards was the result of a report by a Governor's Commission on School Performance. The Commission reported that the state lacked an accountability system that could produce good information on how students in Maryland were doing and who should be accountable for producing high performance by each student within the education system. The Commission, which had broad representation from stakeholder groups, deliberated from 1987 to 1989. As a result of the State Board of Education's acceptance of the Commission's recommendations between 1989 and 1990, representatives of numerous groups from across the state (e.g., teachers, content supervisors, administrators, citizens, parents, special educators, and higher education personnel) worked to reach consensus on performance areas for which schools should be held accountable. From an original list of some 200 variables, the list was narrowed to approximately twelve, with an additional seven other areas reported as supporting information. The twelve variables were linked to state standards, and the seven were used to provide information regarding the difficulty of the education task for a particular building. All of the effort was focused at the building level with companion reports at the district and state levels.

Standards and Curriculum Frameworks in Maryland

Maryland established content standards (learning outcomes) in mathematics, reading, writing, language usage, social studies, and science. The expression of these standards is expected to differ in complexity, language, abstractness, and the structure used to elicit the skill at different grade levels. These content standards were followed by the development of proficiency levels and a single state standard for satisfactory and excellent performance. The content and performance standards set the parameters for assessment practices that would be used in Maryland to hold schools accountable. Performance assessments are conducted to ascertain the extent to which students are meeting or making progress toward meeting state standards, which were established for the year 2000. Schools are responsible for showing continuous improvement toward the standards annually, the publishing of an annual report to the public on that progress, and for involving the business and citizen community in the development of school improvement plans to ensure continuous progress.

The mathematics standards, following closely the standards from the National Council of Teachers of Mathematics (NCTM, 1989), set goals for students in the areas of reasoning, problem solving, communication in mathematics, and understanding mathematics connections within the discipline, with other content areas, and with the real world.

The standards for reading are concerned with the ways students construct, extend, and examine meaning by orienting themselves to various texts and purposes, and the various ways that students interact with texts.

The writing and language usage standards examine the ways students create meaning for various purposes and audiences, using both rhetorical devices and written expression. According to the standards, the purposes served by writing are focused on writing to inform, persuade, and express personal ideas.

The social studies standards have been informed by a number of different sources, including the Maryland Social Studies Task Force, the National Council of Social Studies (NCSS), the National Assessment of Educational Progress (NAEP), the Bradley Commission (named after Maryland's governor), the National Governors' Association, and other significant groups that have been investigating the role of social studies in schools. The standards focus on the knowledge base necessary for understanding history and the social sciences, analysis and application of the knowledge base, and the attitudes necessary to use the knowledge and skills within a context of justice and democratic decision making.

Finally, in the area of science, the standards are based on the principle that science is not a collection of facts nor a collection of processes, but rather that science involves construction of meaning through particular ways of observing, thinking, experimenting, and validating.

In addition to these content standards, the Maryland State Board of Education adopted the Dimensions of Thinking presented by the Association for Curriculum Development (ASCD) as a statewide definition of thinking concepts and skills that must be infused in the assessments. Teams of experts in Dimensions periodically review Maryland assessment tasks to ensure that the concepts and skills are covered in both depth and breadth.

Purposes

The Maryland School Performance Program is used for school improvement and accountability purposes. It includes both student accountability and program accountability. Student accountability is based on performance on the Maryland Functional Testing Program (MFTP). The test is used for high stakes purposes: high school graduation. This program will soon be replaced by a series of 10 end-of-course content assessments that are related to the number and content of graduation requirements. Passage of these state tests will also be a condition of graduation for any student receiving a regular high school diploma.

Program accountability is based on information obtained from the Maryland School Performance Program performance assessments, plus the performance indices and "supporting information" described below.

Components of the Maryland Assessment System

The accountability system in Maryland is called the Maryland School Performance Program (MSPP) and it includes data from two sources: student performance and supporting information. The data on student performance are collected from four sources. First, the Maryland School Performance Assessment Program (MSPAP), which is a totally performance-based assessment given to all students in grades 3, 5, and 8. The assessments do not contain items, but tasks, which may call for group interaction. The results of the assessments produce scale scores in reading, mathematics, writing, science, social studies, and language usage. These scale scores align with five proficiency levels. Each proficiency level describes what a student at the level is able to do. Proficiency level three has been adopted by the State Board as the standard for "satisfactory" performance.

Three additional indices of performance, student attainment (promotion rate and high school graduation rate), student participation (attendance and dropout rates), and postsecondary plans and decisions (grade 12 documented decisions) are also reported annually. All standards are within grade level and school, and data are disaggregated by race and gender.

The second component, supporting information, includes: information on student characteristics (e.g., enrollment and mobility), kindergarten completion, and number of students receiving special services. Other factors reported include financial information, staffing ratios, instructional time, and results of performance on the Comprehensive Test of Basic Skills/4, which is given to a sample of students in grades 3, 5, and 8 in each district and reported only at the district level.

What Do the Tests Look Like?

The Maryland Functional Testing Program includes four basic minimum competency tests: three multiple choice tests in reading, mathematics and citizenship, and a modified-holistically scored direct writing assessment of both narrative and explanatory writing skills. All of these tests are untimed. The multiple choice tests take approximately one and one half hours apiece to give; the writing measure takes approximately two hours, and is given over two days. There are computer-adaptive versions of the reading and math tests, which usually take approximately 30 minutes. The Maryland Functional Testing Program began as a graduation requirement program, and is now required to be completed by the end of eighth grade.

The Maryland School Performance Assessment Program (MSPAP), currently available at the elementary and middle school levels, measures higher order thinking processes and the application of knowledge and skills to real world situations. It is a single test covering mathematics, reading, writing, science, language usage, and social studies. The MSPAP takes nine hours of testing time and is composed of performance tasks with open-ended, short answer, and extended performance based on the State Board adopted outcomes. Item specifications and scoring rubrics for the MSPAP were built by 350 teachers who worked with the Maryland State Department of Education personnel and CTB Macmillan/McGraw Hill. The performance assessments are scored each summer by approximately 650 teachers who are hired and trained by the Maryland State Department of Education personnel and a scoring contractor, currently Measurement Incorporated of Durham, North Carolina (samples of released items and student performances can be found in Appendix A).

Psychometric Properties of the Tests

For each administration of the Maryland School Performance Program a technical report is completed. Similar reports are developed for the administrations of the Maryland Functional Testing Program as well. Maryland made an early commitment, not only to performance assessment, but also to ensuring the psychometric integrity of those assessments. The state's performance assessment for students with severe disabilities, the Independence Mastery Assessment Program (IMAP), is currently being piloted and will also be expected to meet psychometric standards.

In addition to data on the validity and internal consistency of the performance tasks, the scoring contractor is required to provide scorer qualifying information for all scorers before they begin to score tests. Additionally, daily check sets and a daily report on scorer reliability must be produced, and retraining provided for scorers who fall below an agreed upon standard. Approximately five to six scorers are involved in the scoring of an individual student answer booklet.

The Maryland School Performance Assessment Program is divided into three equivalent clusters at each grade level, 3, 5, and 8. Each cluster includes the six content areas, and each student takes one cluster. Assignments to clusters is by a randomization formula provided by the State Department of Education to each of the schools. Each school takes all three clusters. School program data are derived by aggregating across the three clusters. Scale scores are reported at the global content area level (e.g., mathematics), but also at suboutcome levels (e.g., reasoning) to facilitate use by classroom teachers for instructional improvement. Scale scores are then reported at the building, system, and state levels in the context of the five proficiency levels. The sampling design serves to limit the amount of instructional time that is devoted to state testing, while at the same time producing valid and reliable scores at the program level.

What Do the Scores Look Like?

The Maryland Functional Test scores are Pass-Fail. The Maryland Functional Test results are included in the school accountability system by assigning ratings of excellent, satisfactory, or "not met" to the percentage of students passing the tests by the end of grade 9 and the end of grade 11. For the Maryland School Performance Assessment Program three categories are reported related to standards performance: excellent, satisfactory, or not met. They are defined as follows:

- Excellent: a highly challenging and exemplary level of achievement indicating outstanding accomplishment in meeting the needs of students

- Satisfactory: a realistic and rigorous level of achievement indicating proficiency in meeting the needs of students

- Not Met: a level of achievement indicating that more work is needed to attain proficiency in meeting the needs of students

For other standards of pupil performance (promotion rates, attendance rates, dropout rates, etc.) schools are rated on the same levels of state standard attainment of excellent, satisfactory, and not met. The state standards for "satisfactory" govern all schools regardless of demographics or difficulty of the education task. The standards for "excellent" require that a school meet the satisfactory standards and have at least 25% of its students performing at the higher levels.

What Do Reporting Practices Look Like?

The Maryland School Performance Report: School System and Schools is published in December every year for the State and each school system. It includes summary and disaggregated data at the State and school system levels. The numbers of students exempted from assessments are reported. Disaggregated data are reported by gender and race/ethnicity for percent passing the Maryland Functional Test, ratings on each Maryland School Performance Test, and ratings on each performance index (attainment, participation, and postsecondary plans). A comparable report that provides the same data for each school building must be produced by each school system. Each building is responsible for producing and disseminating its own report within the local community. Both system and building reports may go beyond the variables reported by the state, but must use a parallel format. Approximately 16 of 24 systems have exercised the option to add local variables that are of concern to the immediate community. Examples of these additional variables are advanced placement tests, elementary parent-teacher conferences, and number of volunteers per school.

How Are the Data Used?

Each school that is not meeting the state standards must have a School Improvement Team. School level reports are provided to each building through the local assessment coordinator. The results are used by the team to develop a School Improvement Plan. The team looks at the

performance of regular and special students by curriculum area, compares actual performance relative to the performance standards, establishes milestones for improvement, and monitors changes in performance at various checkpoints during the school year. The information is used to guide and improve a school's instructional and organizational activities.

The .State Department of Education monitors progress of each school annually under an accreditation provision, known as reconstitution. This provision requires that a school that is not meeting standards must make progress toward those standards. No growth or movement in a downward direction triggers a reconstitution review, which ultimately could lead to a state takeover of the school. Howevér, low performing schools first get technical assistance and additional funding to alter their performance.

The governor introduced budget provisions in the 1996 legislature to provide a recognition program for high performing schools. The final intent is to allow like-schools that are successful to help those that are not.

To What Extent Do Students with Disabilities Participate in the Accountability System and the Assessments?

MSPP requires that all students be included in the accountability process at the designated grades in all data areas unless those students meet exemption requirements that release them from specific areas of participation.

For students with disabilities, each student's Admission, Review, and Dismissal (ARD) school level committee determines, on an individual basis, whether the student meets the exemption criteria. Students, including students with disabilities, may be exempt from the Maryland School Performance Assessment Program if they are:

a) second semester senior transfers from out of state,
b) first time LEP student administration, or
c) not pursuing the Maryland Learning Outcomes (Reading, Writing, Language Usage, Mathematics, Science, and Social Studies).

An alternative set of outcomes, more life-skills oriented, has been identified and developed by the state under the direction of the IMAP Advisory Committee. The Independence Mastery Assessment Program (IMAP) is a pilot in approximately one third of the local school systems in Maryland. In those districts, the students with severe disabilities participate in either

MSPAP or IMAP, generally depending on which outcomes they are pursuing. The various ways the students with disabilities in non-pilot districts participate in MSPAP include:

a) total participation in MSPAP with no accommodations,
b) total participation in MSPAP with accommodations, and
c) total exemption from MSPAP for students who meet criteria for exemptions.

The various ways the MSPAP/IMAP pilot districts participate include:

a) same as a above,
b) same as b above,
c) total participation in IMAP. IMAP students have severe cognitive disabilities that prevent them from successfully completing regular education course work even with accommodations. This does not mean, however, that they are not included in the regular education program or that they cannot participate in regular school classes and/or activities.

There is flexibility built into MSPAP/IMAP decisions to allow these students to move, when appropriately determined by the school ARD committee, between MSPAP and IMAP. The three components of IMAP include:

a) performance tasks,
b) parent survey, and
c) a student portfolio.

Students with disabilities are included in all other data elements of the Maryland School Performance Program (MSPP) including the Maryland Functional Testing program, a graduation requirement.

What Accommodations Are Permitted?

Accommodations are permitted in the administration of the various assessments of MSPP (MSPAP, MFT, CTBS/4, and IMAP). The state has produced guidelines to assist in the determination of appropriate accommodations. Some general principles include:

- All students are to be included to the fullest extent possible in all statewide assessment programs.

- Decisions regarding exemptions should be made in the ARD/IEP meeting and documented in the IEP.

- It may be appropriate to excuse a student from a state assessment program administration. The Local Accountability Coordinator is directly responsible for making final decisions and clarifications about exemptions, but in collaboration with Special Education, LEP, 504, and school-based staff.

- Single or multiple accommodations that may be difficult to provide are **not** a reason to exempt a student.

- Students may be exempted from an individual MSPAP content area as a result of
  (1) their instructional program, or
  (2) specified accommodations.

The accommodations should reflect the student's ongoing instructional program including classroom assessments.

For each of the state assessments, accommodations may include changes in: scheduling, setting, equipment, presentation, and response. These five areas are further delineated for each statewide assessment (i.e., MSPAP, MFT, CTBS/4).

Implementation Considerations

The implementation of Maryland's assessment and accountability system is broad, requiring personnel, resources and costs, training, and monitoring.

Personnel. MSPAP State Level–The Maryland State Department of Education operates the design, development, and implementation of its three state level testing programs through the Assessment Branch of the Division of Planning, Results, and Information Management (PRIM). That branch has a total of 10 staff members. Additionally, for the Maryland School Performance Assessment Program (the state's performance assessment) curriculum specialists in language arts, reading, mathematics, social studies, and science are assigned to work collaboratively on the development of the assessments. Their time commitment is approximately 75%. Reports on the results of the assessments are carried out through the Results branch of the PRIM. It should be noted that in 1987 the Assessment Branch consisted of approximately 12 staff members. Hence, the performance assessment program has been

implemented through cooperative activities with the curriculum specialists with virtually no addition in staff in the area of assessment.

MSPAP Local Level—At the local school system level each system has a formally appointed Local Accountability Coordinator. This individual, named by the local superintendent, is directly responsible for linkages to the State Department of Education. Additionally, each building has designated a school test coordinator, who is responsible for the administration and management of the assessments, particularly the Maryland School Performance Program, at the building level.

To begin the development of the performance assessment in grades 3, 5, and 8*, each school system was invited to send 15 teachers (five at each level) to assist in the development. These individuals were released approximately 20 days during the school year. Additionally, approximately 12 curriculum supervisors worked approximately 30 days in the development effort. This major commitment of local time was supported by the local superintendents as crucial to local teachers' understanding the assessment and ownership of the product.

IMAP State Level—The planning and development of alternate outcomes and assessment procedures was initiated by the IMAP Advisory Committee with the assistance of a federally funded grant to develop state outcomes and indicators for this group of students with severe disabilities. One staff member at the state spent approximately 30% time on the project. The advisory committee was identified and selected to serve, with little or no additional cost, primarily through participation in meetings. After a series of initial meetings, the federal grant was secured to assist the development process and to focus the development of the outcomes and assessment system. The federal funding greatly accelerated the development process and relieved the state from some initial start-up costs for this assessment. Once begun, allocations for a core group of teachers, substitutes, and summer workshops were developed. Assessment procedures and process were designed, initiated, and have continued to be developed and refined after the end of the federal grant.

In Maryland, ten to fifteen stakeholders, one SEA staff, one to two Institutions of Higher Education (IHE) personnel and eight to ten teachers working one week in the summer were sufficient to begin the development of IMAP assessments. Teachers were added each summer to serve as trainers in their school systems the following year. Each year, a grade/age group was added to be trained and to assist in the scoring process.

* Maryland has a proposed high school assessment that will be developed and phased-in for administration in grade 8 in the near future.

General Cost Factors. Although we do not provide a cost analysis, there are several cost factors that can be shared. In the interest of generalization to other states, this discussion of cost focuses at the cost per pupil level since that cost is not as readily influenced by the size of the state or the number of school districts.

Regular Assessment System Costs—The state's norm-referenced program (CTBS) typically operates at approximately $5-6 per student. The Maryland Functional Testing Program, including tests in reading, mathematics, citizenship, and the direct assessment of writing, operates at a cost of approximately $12-15 per student. The Maryland School Performance Program was designed, developed, and implemented at a cost of approximately $21-23 per student. While this cost may be considered high compared to traditional norm-referenced selected-response testing, when viewed from the perspective of quality control, this cost is a minor part of the $5,000-8,000 per pupil being spent within public education today.

These estimates do not include the cost of State Department of Education assessment staff, since no new staff were added as a result of the programs. Indeed, if any of the programs were to cease to exist, there is no assumption that staff savings would accrue. Hence, to include their costs would appear to inflate the figure in a way that would not represent the true costs of the program.

IMAP Alternate Assessment Costs—The cost of the alternate assessment in Maryland has been relatively nominal. Data on costs are shown in Table 1. Performance assessments required paying teacher stipends to develop an initial bank of performance items. Parent surveys were approximately $400 per local school system for a site license to make copies. The cost of portfolio training and development were nominal and completed by teachers during the summer workshops. Videotape, used for recording the individual student performance tasks, was purchased in bulk, costing approximately $3.50 per VHS tape.

All of the these costs are developmental and linked to pilot testing, since the program has yet to be implemented for all students eligible for IMAP. As that time approaches, a more comprehensive implementation cost analysis will be conducted.

Training Issues. Regular Assessment System Training—As mentioned earlier, 15 teachers from each school system were invited to participate in the development of the assessment. Annually, task development for new editions of the tests take place using approximately 100 teachers. All of these individuals received training in task development and test specifications prior to task development.

Additionally, the MSPAP is scored each summer by approximately 600-650 Maryland teachers. Each of these individuals is trained in the scoring methodology and monitored on a daily basis for the maintenance of scorer reliability. Daily reports are submitted by the scoring contractor, and remedial training occurs for individuals whose scoring reliability falls below an acceptable level. All scorers are hired by the contractor, so that scorers unable to reach or maintain an acceptable level are terminated.

IMAP Training—Training in the alternate assessment in Maryland's IMAP takes place in two stages. First, a presentation is given in the local school system for awareness and to respond to questions. The second phase is a two week training session in the summer with one week for scoring and the second week for development of new performance tasks and revision of old tasks. Every year new teachers are trained and the process is reviewed. There are two review panels: one of field experts to review the content, appropriateness, and completeness of the outcomes, indicators, and rationale; the second panel reviews the technical soundness of the assessment process and the congruence with the other state assessments.

Monitoring. Regular Assessment System—Scores are monitored for unusual peaks and valleys annually. In fact, the monitoring system has identified several instances of cheating. The State Board of Education has established a stringent policy regarding assessment irregularities, and several teachers have had their licenses suspended or revoked permanently.

IMAP Alternate Assessment System—IMAP monitoring currently is a local school system process with the appointment of a program coordinator/director and accountability coordinator to carry out the duties of monitoring. This is expected to change as the program moves from a development to an implementation phase.

Table 1. Estimated Cost of Initial IMAP

The estimated cost of IMAP development is based on cost summaries and factors included in the development of the system. As changes emerge and refinements are made, cost factors will vary. It is important to understand that start up costs generally are higher than maintenance costs. The cost estimates are based on per student estimates.

ITEM	COST	NOTES
Activity Materials	$6.00*	Vary according to task
Parent Survey	$3.00*	Site license for entire LSS $400
Video Tapes	$3.50*	Bulk rate when using MPT possible
Task Scoring	$18.75*	One teacher scoring eight videos per day @ $150
Task Development Revisions Outcomes/Indicators	$75.00†	One teacher develops two tasks per day @ $150
Portfolio a) Basic b) Optional	 $5.00†† $30.00	 Basic components and packaging CD ROM site (LSS) with SEA assistance
Substitutes	$23.00	Training assumes 2-3 students per teacher/sub. Pay @ $70 per day for one day training
Staff Development	$27.00	Trainer of trainers model/$2.00 per in-service, 75 students per LSS

Note: All costs are based on an estimate of 1,803 students in the state and 75 students in each local school system (LSS). Totals are:

$191.25 per child for initial start-up cost ($14,344 per LSS)

$31.25 per child for basic cost and ongoing costs ($2,343 per LSS)

* Each year cost for minimum per student assessment amount

† Reduced each year depending on amount of new tasks required and revisions necessary

†† Basic amount, generally a one time cost

Assessment Practices in Kentucky

Overview

The Kentucky Education Reform Act (KERA) of 1990 formed the basis for massive change in the state's educational system. This massive reform was enacted by the Kentucky General Assembly as a result of a lawsuit brought against the General Assembly, the governor, the Kentucky Board of Education, and the chief state school officer by the Coalition for Better Education (CBE), which represented approximately 60 of the state's 176 school districts. The successful 1988 lawsuit found the state's funding mechanisms inequitable and mandated that the educational system be redesigned. One of the most comprehensive, statewide restructuring efforts ever attempted in the United States, the reform called for top-down and bottom-up systemic change in finance, governance, curriculum, and assessment.

KERA established six goals for the schools of the Commonwealth: (1) expect a high level of achievement of all students; (2) develop students' abilities in six cognitive areas; (3) increase school attendance rates; (4) reduce dropout and retention rates; (5) reduce physical and mental health barriers to learning; and (6) increase the proportion of students who make a successful transition to work, postsecondary education, and the military.

The first and second of these goals were translated into a "primarily performance-based assessment" program using two approaches. First, a Council on School Performance Standards was created to further define the six cognitive areas related to student ability. Through a two-year period of public input and review, the Council arrived at 75 performance goals now known as "Academic Expectations." Through various review processes, these expectations were refined and consolidated; 58 are assessed formally within the measurement system. The Council presented the results of its efforts to the State Board of Education in December, 1991.

The other major activity was to establish an assessment system, the Kentucky Instructional Results Information System (KIRIS), to measure progress toward the goals, primarily the academic expectations reflected in the first two goals of the act. Through a competitive process, the Kentucky Department of Education selected Advanced Systems in Measurement and Evaluation as the contractor for the assessment program. In addition to creating and operating the program based on the Academic Expectations, Advanced Systems was required to assist the Department in creating operational definitions of the third through sixth goals, referred to as the "non-cognitive goals." Advanced Systems hired a subcontractor, the Far West Laboratory for Educational Research and Development, to develop recommendations for the non-cognitive goals for the Department.

Standards and Curriculum Frameworks in Kentucky

Kentucky's reform effort is based on the notion that all students are capable of learning. The expectations for students outlined in the six learning goals and the identified academic expectations provide the frameworks for the high expectations of all students.

The six Learning Goals of the KERA, as outlined in the second of the six major goals of the act, are as follows:

1. Students are able to use basic communication and mathematics skills for purposes and situations they will encounter throughout their lives.

2. Students shall develop their abilities to apply core concepts and principles from mathematics, the sciences, the arts, the humanities, social studies, practical living studies, and vocational studies to what they will encounter throughout their lives.

3. Students shall develop their abilities to become self-sufficient individuals.

4. Students shall develop their abilities to become responsible members of a family, work group, or community, including demonstrating effectiveness in community service.

5. Students shall develop their abilities to think and solve problems in school situations and in a variety of situations they will encounter in life.

6. Students shall develop their abilities to connect and integrate experiences and new knowledge from all subject matter fields with what they have previously learned and to build on past learning experiences to acquire new information through various media sources.

The Alternate Portfolio Advisory Committee, a committee consisting mainly of teachers of students with moderate to severe disabilities, identified a subset of Kentucky's Academic Expectations for use in the Alternate Portfolio Process. There were 28 expectations identified as critical to maintaining the integrity of functional programming for students participating in the Alternate Portfolio Process (p.6, KIRIS Kentucky Alternate Portfolio Project–Teacher's Guide). This subset of Kentucky's Academic Expectations is used to assess the performance of students participating in the Alternate Portfolio system. The Alternate Portfolio is intended to be used with that small percentage of students who cannot participate in the mainstream

assessment process, even with the assistance and/or adaptations that can be made available to the instructional process.

Purposes

The purposes of the KIRIS assessment system include:

- focusing instruction on the need to develop means by which all students can learn and demonstrate learning at high levels;
- focusing instruction on the application as well as the acquisition of skills;
- providing incentives for school staff who make significant progress in improving instructional services as evidenced by significantly higher performance of students on the KIRIS assessment components; and
- providing assistance to school staff who indicate the need for assistance in bringing about these desired improvements as evidenced by student performance on the KIRIS assessment components.

Components of the Kentucky Assessment System

The contents of the KIRIS assessment components are influenced primarily by the direction of content area advisory committees with members drawn mostly from classrooms, schools, professional education organizations, higher education, community groups, the Department of Education, and Advanced Systems in Measurement and Evaluation. The KIRIS assessment, which has been administered annually from 1991 to 1996, includes three types of assessment tasks:

> Assessment tasks involving portfolios—Each student in grades 4, 8, and 12 is required to assemble a Writing Portfolio and a Mathematics Portfolio (as of the 1994-95 school year Mathematics Portfolios are required in grade 5, rather than grade 4). These portfolios represent collections of the student's best work developed over time in conjunction with support from teachers, peers, and parents. The portfolios are scored by local teachers, and the scores are reported to the Kentucky Department of Education for use in the accountability assessment. Mathematics portfolios will not be included in the baseline

calculation for 1996-97 and 1997-98, but will be included for instructional purposes in 1997-98, and for accountability purposes in 1998-99.

Assessment tasks involving performance events—Students participate in performance-based assessment tasks that require them to use knowledge and skills learned in school to produce a product or solve a problem. Rather than recall facts, students apply what they have learned to a real (or real-life simulated) situation. Performance event tasks involve both group and individual work, are based on manipulatives and/or other materials, and take about an hour each for completion. Performance event tasks are administered by test administrators hired by Advanced Systems in Measurement and Evaluation. For 1996-97 and beyond, performance events enter a research and development phase. Until this is complete, they will not be included in the accountability index.

Assessment tasks involving open-ended questions—Students respond to open-ended questions requiring extended written responses. The focus is on higher-order thinking skills, solving multi-step problems, and using reasoning, analytical, and written communication skills.

Assessment tasks involving machine-scorable questions—In 1992-94 students also answered a section of multiple choice questions, although these were not used for accountability purposes. Beginning in 1994-95, KIRIS included a section of other item types being evaluated for possible inclusion in the future. Beginning in 1996-97, a section of multiple choice questions will be included in each content area for accountability purposes.

KIRIS also monitors school progress in terms of non-cognitive indicators such as school attendance rates, dropout and retention rates, reduction in physical and mental health barriers to learning, and the proportion of students who make a successful transition to work, postsecondary education, or the military.

Within the accountability system used in Kentucky, school performance is reported as an overall Index score, an Index score for each of the cognitive areas, and for the non-cognitive measures taken together. Students' results in the cognitive areas are reported as four performance levels: Novice, Apprentice, Proficient, and Distinguished.

The Alternate Portfolio Assessment process is a multi-disciplinary approach as opposed to a single curriculum area. It follows the model of the Kentucky mathematics and writing

portfolios in using a holistic scoring guide. The following are key concepts of the Alternate Portfolio:

- Scores of students participating in the assessment are weighted equally with those of students participating in the regular assessment for the school's accountability purposes.

- Entries to the student's portfolio are not specified, other than that each entry must be related to the state's Academic Expectations.

- An Alternate Portfolio Advisory Committee, charged with the task of identifying the Academic Expectations to be assessed within the Alternate Portfolio process, first looked at the critical functions of each of the Academic Expectations and determined the extent to which each could be evidenced for children eligible to participate in the Alternate Portfolio process. Through a process of reconciling student activity schedules and evidenced performance, the Committee initially identified 28 academic expectations critical to maintaining the integrity of functional programming for students participating in the Alternate Portfolio process. Examples of the critical function for each outcome have also been identified. It is expected that other Academic Expectations will be incorporated during subsequent years. (The KIRIS Alternate Portfolio Assessment was implemented for the first time during the 1992-93 school year; the first set of scored portfolios were completed in the fall of 1993. Because the Alternate Portfolio program did not begin until the 1992-93 school year, students in this category were exempted from testing for 1991-92.)

What Do the Tests Look Like?

The KIRIS Transitional (On-Demand) Assessment currently consists of challenging open-response items. A student may be asked to read a real selection from grade-appropriate literature and demonstrate comprehension of that passage by writing several paragraphs demonstrating such understanding. KIRIS Performance Events require students to complete group activities first and then to produce an individual student written response to a specific question or questions. For example, students may be asked to perform a brief science experiment and record the resulting data in a group setting before individually responding to an open-response item (samples of released items and student performances can be found in Appendix B).

Portfolio entries ideally should resemble normal classroom work and should represent the student's best work in that content area, determined mostly by the student with advice from the instructional staff.

Psychometric Properties of the Various Measures

Content validity. The KIRIS assessment components address content validity in a traditional manner. First, the Kentucky State Board of Education established a set of academic expectations that were to form the basis for the KIRIS assessment process. Content Advisory Committees were established in the subject areas of reading, mathematics, science, social studies, writing, arts and humanities, and practical living/vocational studies. In addition, an Alternate Portfolio advisory committee was created for the assessment mechanism that crosses content area lines and is designed for student with severe disabilities. These committees consisted mostly of teachers with expertise in the content areas tested. The content advisory committees reviewed assessment items and tasks to assure that they did address the academic expectations and identified core content.

Consequential validity. In an environment where assessment results have significant impact on the instructional process, the question of consequential validity must also be considered. The components of the KIRIS assessment were designed to encourage certain kinds of instructional practices. While addressing traditional instruction in the basic skills was encouraged, the KIRIS assessment components clearly required instruction to consider the student's ability to apply both basic skills and higher-order kinds of skills in the form of constructing responses to rather complex and challenging assessment tasks.

Reliability and generalizability. Although accountability decisions take place only at the school level, results are reported to parents at the student level. It is useful, therefore, to examine reliabilities at that level. In Table 2, we list student-level coefficient alphas for open-response items for the first accountability cycle. These figures are calculated using results from students who were eligible to complete open-response tests and who were present on the day of a testing, with absence being determined by having blank responses to all items; blank responses by students who had attempted at least one question were counted as zero in the computation of coefficient alpha. The reason for not including the responses of absent students is that score reliability would be slightly overestimated.

Table 2. Open-Response Test Reliabilities by Subject and Grade

	Subject	1991-92 Open-Response (three items, 12 possible score points)	1992-93 Open-Response (five items, 20 possible score points)	1993-94 Open-Response (five items, 20 possible score points)
4	Reading	.58	.77	.79
	Mathematics	.63	.72	.75
	Science	.63	.65	.73
	Social Studies	.64	.69	.76
	Composite	.88	.92	.93
8	Reading	.72	.83	.80
	Mathematics	.61	.79	.71
	Science	.62	.75	.66
	Social Studies	.74	.81	.83
	Composite	.89	.94	.92
12	Reading	.79	.85	.75
	Mathematics	.65	.81	.76
	Science	.72	.81	.76
	Social Studies	.77	.85	.85
	Composite	.92	.95	.94

NOTE: The above information is from pages 196-198 of the KIRIS Accountability Cycle 1 Technical Manual, KDE, July 10, 1995. The reliabilities were computed by form for common and matrix items, and averaged.

With the exception of grade 4 Science, open-response reliabilities increased from 1991-92 to 1992-93, a logical result of increasing the number of items from four to seven. The 1994-95 assessment includes an additional item in grades 8 and 11†.

As a caveat, it should be noted that using coefficient alpha probably underestimates score reliability insofar as item raw scores are the basis for the computation, whereas the fundamental

† Although the Kentucky Department of Education advises against making student-level decisions based on individual test scores alone, open-response test reliabilities compare favorably with reliabilities from other tests used to make student-level decisions. Composite score reliability is comparable to reliabilities noted for the ACT Composite, and individual subject area reliabilities are similar to ACT and CTBS subject area reliabilities, as well as to reliabilities for a national certification test in accountancy.

scaling method employs a logistic model. The use of item response theory takes into account differences in item difficulty not reflected in the computational use of raw scores. A five item test having a single, relatively difficult item is likely to yield a diminished coefficient alpha in comparison with results from a comparable test with essentially uniform item difficulty.

Student-level reliability estimates (in terms of scorer agreement) for portfolios and performance events are addressed in Kentucky's technical reports. Given that each student produces only one writing portfolio and, with limited exceptions, takes only one performance event, score reliability cannot be examined in the same manner as open-response items.

Who Takes the Tests and What is the Unit of Analysis?

In general, students with disabilities take all of the regular KIRIS assessment components, either without adaptations as do all other students, or they participate in the regular components of the KIRIS assessment with that assistance and/or adaptation normally made through the daily delivery of instructional services. A small percentage (generally less than one percent of the total accountability grade) of the population of students with disabilities participate in the KIRIS Alternate Portfolio Assessment; if the student with disabilities, with the aid of all assistance and adaptive devices that can be made available, cannot participate in the regular curriculum, then the student participates in the Alternate Portfolio process.

What Do Reporting Practices Look Like?

Each school district is required to publish an annual report that informs the public about its performance in each of the critical areas used to determine school success. In addition, a system for determining successful schools has been established in Kentucky based upon student achievement relative to both the school's improvement goals and the state goals.

How are the Data Used?

In January 1993, using the results from the spring 1992 administration of the KIRIS, the Kentucky Department of Education assigned to each of Kentucky's approximately 1,255 schools (1) a baseline score on a metric referred to as an accountability index, and (2) an improvement goal (previously referred to as a threshold), which is a two-year target for improvement. The accountability index in the first cycle of implementation (1991-92 through 1993-94) was a number ranging from one to approximately 133 that is based on six

components: assessment scores for the five content areas of reading, writing, mathematics, science, and social studies; and a number representing a score for the combined noncognitive indicators. The accountability system is not intended to compare schools or districts to each other; instead, it provides both a school baseline score and a target for improvement for each school and district. Every two years each school will receive a new baseline accountability index and a new improvement goal based upon its students' performance on the assessment. This means that the school must continue to improve student achievement, as measured by KIRIS, every two years.

This same information applies to students with disabilities. Their results include performance results on the KIRIS Alternate Portfolio Assessment and data for the non-cognitive indicators; both are included in determining a school's accountability index. Data reported for schools and school districts may be disaggregated for students with disabilities by disability category at grades at the elementary, middle, and high school levels.

Schools receive rewards and assistance (sanctions) based on their performance. Schools achieving above their improvement goal (threshold score) receive financial rewards. Schools failing to achieve their target are required to develop school improvement plans with assistance from the Kentucky Department of Education. KERA requires that a school scoring more than five percent below the school baseline be declared a "school in crisis." After learning that the school is in crisis, parents may transfer their children to other schools. To assist the school considered to be in crisis, the Department of Education will assign a "distinguished educator" with statutory powers to enforce sweeping changes, including staff assignments. Implementation of the "school in crisis" provision of KERA was delayed by the Legislature until at least 1996. If an entire district fails to meet academic goals or make significant improvement, the commissioner and state board can remove the local superintendent and board members from office and appoint replacements.

Results of the first accountability biennium announced in January 1995 indicated that 95% of Kentucky schools had improved over their baseline, with almost one third eligible for and receiving financial rewards.

To What Extent Do Students with Disabilities Participate in the Accountability System and the Various Assessments?

The Kentucky Education Reform Act of 1990 requires the inclusion of all students in the accountability process at the designated grades. For students with disabilities, each student's Admissions and Release Committee determines on an individual basis how the student will be

included in the KIRIS assessment program. Students with disabilities must participate in the KIRIS assessments in one of the following ways:

- Full participation in all three components of the assessment program with no adaptations or modifications;

- Full participation in all three components with adaptations and modifications, including the use of assistive technology devices that are consistent with the instructional strategies specified on the student's IEP or 504 Plan and available to the student in the course of his/her instructional process;

- Participation in an Alternate Portfolio Assessment program. Students who meet the established criteria participate in the Alternate Portfolio Assessment program. These students are generally those who have moderate to severe cognitive disabilities that prevent them from completing a regular course of study even with program modifications.

Students participating in the Alternate Portfolio assessment do not participate in the other components of the assessment [i.e., Transitional (On-demand), Performance Events, writing portfolios and mathematics portfolios]. Schools, as well as the state as a whole, are expected to place no more than two percent of their population in the Alternate Portfolio Program. Schools that do exceed this percentage would be monitored by the state agency; however, no school has exceeded this criterion.

What Accommodations Are Permitted?

Accommodations may include changes in the administration of the assessment and/or recording of student responses that are consistent with the normal instructional strategies and assistive devices and services identified on the student's IEP or 504 plan. Adaptations in the instructional process must be related to the student's disability and specially designed instruction as described on the student's IEP or 504 plan, and must be age-appropriate. It is also specified that adaptations shall not inappropriately impact the content being measured. Reading assessments may be read to a student if the intent of the assessment is to measure comprehension, but only if this is the normal mode through which the student is presented regular print materials and if this is documented on the student's IEP or 504 Plan.

What Are the Implementation Considerations?

The implementation of Kentucky's accountability system requires personnel, resources and costs, training, and monitoring.

Personnel. The need for additional personnel to implement policies related to the participation of students with disabilities in the regular KIRIS student assessment components with adaptations and assistance is certainly difficult to project above and beyond the normal personnel needs of the regular instructional program. In theory, there would be no additional needs because these kinds of staff would be needed throughout the instructional year and would therefore be available during the KIRIS assessment administration. In practice, local school districts may have to reassign staff, seek volunteers, etc., in order to meet the simultaneous needs of students. Under normal instructional workloads, local schools may find it easier to schedule special assistance so as not to require the numbers of staff required during KIRIS administration. The Alternate Portfolio process does require two to three staff members at the state level to coordinate advisory committee activities and provide training.

Costs. The cost for operating the Alternate Portfolio program from the state perspective is approximately $75,000 to $90,000 per year. However, there are additional uncalculated costs associated with LEAs providing release time, etc., to allow teachers involved in developing and scoring the alternate portfolios to attend the necessary training and scoring activities.

Training related to the scoring of KIRIS transitional (on-demand) assessments. The answer documents of students with disabilities are scored in the same way as those of the total population. In the KIRIS system, prior to training the scorers, the test developers and scoring trainers meet to review the test items, scoring guides, and more than 200 student responses in order to select those to be used in the training pack. Scoring training involves from 210 to 350 scorers divided into smaller groups (30-50) to work with two trainers in each content area. Scorers are trained on a particular test form and then score those responses in a group to keep the scoring guides fresh in their minds. Scores are supervised by training leaders, and two percent are rescored by the supervisor to test reliability. Training leaders rescore a minimum of four papers per scorer per day, to ensure continued reliability.

Monitoring of the implementation of alternate portfolios and KIRIS modifications. To date, student participation in the Alternate Portfolio assessment program has been less than 0.5% of the total student population, with very small increases since its inception. It is estimated that approximately 15% of the student population participate in the KIRIS assessments through the use of modifications each year.

Comparison of Assessment Practices in the Two States

Overview

In both Maryland and Kentucky the assessment and accountability systems came about in response to pressures for accountability. In Maryland the pressure was from the legislature, while in Kentucky it took legal action to drive reform. Maryland and Kentucky are the two states that have made the most progress in including all students, including students with disabilities, in their assessment systems. In this section a comparison of the assessment practices in the two states on a number of dimensions is made.

Definition of Accountability

The Maryland School Performance Program has both student accountability and program accountability components. Student accountability is high stakes for students. They must take and pass the Maryland Functional Testing Program (MFTP) in order to graduate from high school. School accountability is based on performance on the Maryland School Performance Assessment Program (MSPAP). MSPAP and MFTP are both components of Maryland's larger School Performance Program (MSPP). MSPP includes a number of school data elements, including attendance, norm-referenced or criterion referenced assessments, dropout rate, etc.

The Kentucky School and School District Accountability program is primarily an accountability model that hold schools and districts accountable for demonstrating improved delivery of instructional services directed at high academic expectations. The school is held accountable for improving its delivery of instruction until in the ideal sense it meets the long range goal of producing a student body that:

- Is proficient in all academic content areas measured,
- Attends school 100% of the instructional year,
- Is not retained and does not dropout of school, and
- Makes a successful transition to adult life.

While there are certain compensatory features in the Kentucky system, the above description does state the ideal condition under which a Kentucky accountability index of 100 could be achieved. No schools or districts are currently meeting this ideal goal even considering the compensatory nature of the scale. Schools and districts are required to reduce the distance

between their index score and 100 by 10% on a biennial basis. Student accountability is certainly a component that Kentucky recognizes as important, but the current model leaves student accountability to the discretion of local schools and/or districts.

Measurement of Progress/Performance

Progress in Maryland is measured by setting five-year performance goals and then monitoring school system progress toward meeting those long-range goals. In Kentucky progress is measured by establishing baseline performance for each school and then monitoring change in performance relative to the baseline and the long-range state established goal.

Consequences (Rewards and Sanctions) for Student or District Performance

In Maryland there are both student and school accountability systems. The student accountability system has one major consequence: high school graduation. The school accountability system has sanctions rather than rewards, although rewards have been proposed to the 1996 legislature. The state department of education monitors school progress. Schools not meeting standards must make progress toward meeting those standards. No growth triggers reconstitution review and may lead to state takeover.

In Kentucky there are both rewards and sanctions (or assistance provided to schools or districts not meeting biennial expectations) as part of the school accountability system. Schools and/or districts exceeding their improvement goals are to be given financial rewards of which they, as a staff, determine the dispersal. Schools failing to meet their improvement goals but still improving are required to produce a specific school improvement plan addressing how greater progress will be made in the future. After a biennium, schools that are in decline or have scored below their baseline are, in addition to designing the improvement plan described above, also assigned a Kentucky Distinguished Educator who will serve in an advisory role assisting in implementing the school's improvement plan and in monitoring progress. After 1996, schools declining substantially (by five points or more on the Kentucky index) will be considered schools in crisis. The assigned Kentucky Distinguished Educator must basically take over the school. After first placing certified staff on probation, the Distinguished Educator will then notify parents of students served by the school of the school's status and of their options as parents to have their child placed in a school not in crisis, if they so choose. The sanctions become progressively stiffer if the school or district continues to fail to meet its improvement goal.

Rates of Participation of Students with Disabilities

In Maryland, all but about one percent of the students participate in the state accountability system. Most students take the MFTP and the MSPAP. Some students work toward alternative objectives and are assessed using the IMAP.

In Kentucky, all students, including students with disabilities, participate in state assessments and thereby the state school accountability program. Approximately 85 percent of the students take the regular state assessment, approximately 15 percent take the state assessment with accommodations, and 0.5 percent participate in the Alternate Portfolio assessment system.

Accommodation Rules

Both Maryland and Kentucky permit accommodations in assessments. Maryland publishes a set of very specific rules about allowable accommodations. Kentucky students are permitted any accommodations in assessment that are also permitted in and consistent with the appropriate delivery of instructional service.

Is Performance Assessment Necessary to a Good Assessment System?

Performance assessment is critical to any assessment system intended to serve within a high-stakes school accountability environment or in any assessment program that is otherwise intended to influence the quality of the instructional program. Where assessment influences instructional practices (intended or unintended), the assessment model implemented must address the content that the curriculum is to be centered around, the desired modes of instruction, and the level of performance toward which the student must strive. Where the results of an assessment process have intended (or unintended) consequences for a school as a whole, the teaching staff must spend some portion of their instructional time preparing students to perform well on the assessment. Therefore, it is not just a matter of including performance assessment within the assessment process, but these performances must be of the kind that will encourage desired instructional practices within the daily delivery of instructional services, and must be visible enough within the assessment process to draw attention to the desired instructional practices.

The Nature of Participation and Accommodations Decisions

The decisions about which accommodations for students with disabilities can be used in the Kentucky assessment process are individually made with the intent that these decisions permit adaptations consistent with the normal and appropriate delivery of instruction.

A major feature of the assessment systems in both states is that participation and accommodations decisions are individualized rather than categorical (based on disability category) in nature.

Comparison of Practices in Maryland and Kentucky to Other States

States with Standards

These two states are certainly not alone in the enterprise of establishing educational standards. Almost all states are in the midst of such reform. In its 1995 report *Making Standards Matter*, the American Federation of Teachers found that 27 states already had established an identifiable set of educational standards, and 23 other states (including Washington, DC) were in the process of establishing them (AFT, 1995). Iowa remains the only state not focused on establishing academic standards.

Where Kentucky and Maryland hold a comparative advantage over many other states is in their consideration of students with disabilities throughout the various stages of standard-setting. Taking the term *all* to really mean *all*, these two states were able to establish policies and procedures that considered and provided the means by which students with disabilities could be included, from the very earliest stages of the process. And, in developing assessments of the extent to which students are achieving standards, Kentucky and Maryland have been developing assessments for all students.

Participation Decisionmaking

In contrast to Kentucky and Maryland, there is great variability in the rate at which students with disabilities participate in assessments in other states. In its 1994 survey of state special education practices, NCEO found that state directors of special education could report participation rates for only 49 of the 133 tests administered that year, less than 37% of the national total. For those cases where participation rates *were* reported, there appeared to be

wide variability in the degree to which students with disabilities were participating (Erickson, Thurlow, & Thor, 1995).

In analysis of state policies about participation of students with disabilities in assessment, Thurlow, Scott, and Ysseldyke (1995b) suggest that decisions are based most often on one or more of the following criteria:

(1) the primary educational setting of the student (e.g., special school, residential facility, ungraded program, or special classroom);

(2) the student's disability category;

(3) the student's level of access to the general curriculum. In the past, for example, NAEP participation criteria allowed for the exclusion of a student with an IEP if that student had been mainstreamed "less than 50 percent of the time in academic subjects and is judged to be incapable of meaningfully taking part in the assessment" (Mullis, 1990); or

(4) case-by-case decisions of individual administrators or IEP teams. In many cases, this decision must be documented on the student's Individualized Education Program (IEP).

These various eligibility criteria can make the comparison of participation rates between schools, districts or states problematic, if not impossible, since the reported rates may or may not be including many special education students in the population being used as the reference group.

In a very real sense, Kentucky and Maryland eliminate the confusion of such policies by having fully inclusive participation policies in place. For these two states, it is not a question of *who* participates, simply a question of *how*. And that question is given to the people closest to the individual student to decide, along with unambiguous guidelines regarding the use of accommodations, or placement into the alternative testing program.

Participation and High Stakes

In both Kentucky and Maryland, policymakers recognized the fundamental relationship between participation rates and high stakes testing–that under a system of accountability that compares the performance of schools and districts (and often make awards or sanctions based

on these results) administrators and teachers will be motivated to minimize the number of low-performing test takers in order to raise their school or district's overall performance. Zlatos (1994) examined 14 major urban school districts and found that participation rates in testing varied from 93% of all enrolled students in Memphis to 66% in Boston. This system of academic "red shirting" of students with learning problems has resulted in a system that perpetuates invalid comparisons among our nation's schools and school districts.

Both these states have created reporting mechanisms to audit the exclusion of students with disabilities from participation in their large-scale assessment programs. And both have in place alternate assessments designed to integrate the performance data of students with severe disabilities with the scores of other students, thus ensuring accountability for all students being served in their schools.

Using Assessment Information for Improving Programs

Participation by students with disabilities in assessment programs serves many purposes, not the least of which is when the information it gathers can be used for programmatic and policy decision making by either local or state educational agencies. Unfortunately, only a small number of states report using the data in this manner. In its 1995 national survey of the 50 states and 10 special territories, NCEO found that nine state directors of special education did not know whether the performance scores for students with disabilities could be disaggregated from their state's assessment database, and another 20 stated that they definitely could not compile such information. Only 25 states or territories could identify the performance scores of students with disabilities, and out of those, only 12 reported actually analyzing the scores (Erickson, Thurlow, Seyfarth, & Thor, 1996). Because of the inclusiveness of their systems, both Kentucky and Maryland have begun to analyze the achievement data of students with disabilities and have plans to make performance reports available to local and state audiences.

Accommodations in Assessment

Kentucky and Maryland join many other states in their provision of testing accommodations for students with disabilities. In fact, the number of states with formal written policy guidelines on accommodations jumped from 30 in 1991 to 43 in 1995 (Thurlow, Scott, & Ysseldyke, 1995a). That is not to say that their use has not been problematic for state testing officials. Questions continue to surface about the effect of such accommodations on the psychometric integrity of the assessment. With little research to guide them, state assessment offices have

produced guidelines that differ greatly from one another, and often allow testing accommodations to be used that other states specifically prohibit.

Kentucky and Maryland have secured federal funding from the U.S. Office of Special Education Programs (OSEP) for a series of empirical studies focused on the impact of accommodations on performance data. Only two other such grants were awarded from that governmental agency during that particular funding cycle.

An alternative testing program for students with the most severe disabilities, such as Kentucky's Alternate Portfolio or Maryland's IMAP, might be thought of as a major form of accommodation. In this regard, the two states have few equals. As currently stated, Part B of the Individuals with Disabilities Education Act (IDEA) will stipulate the use of alternative testing for those students who cannot participate in a state's regular testing program. This impending change has motivated several states to release requests for assistance in building such alternative systems. State education officials in states such as Texas, Rhode Island and Massachusetts are currently investigating ways to develop an alternate assessment program for students with very severe disabilities. It is anticipated that the lessons learned through their own experiences could make Kentucky and Maryland valuable sources of information as other states undergo such efforts to build truly inclusive accountability systems.

Conclusion

Maryland and Kentucky are two states with inclusive assessment systems. The assessment system in both states came about in response to state legislative pressure for accountability for the results of education for all students. The two states have multiple forms of assessment and explicit procedures for deciding who participates in the various assessments. In both Maryland and Kentucky, schools are held accountable for improved student performance. In Maryland, schools are expected to show progress toward state-defined long-term goals. In Kentucky, schools are expected to improve over baseline performance.

Maryland, Kentucky and most other states permit accommodations in assessments. Yet, Maryland and Kentucky are among the very few states that have an alternate assessment system, one that permits participation by students with severe disabilities. We believe the assessment and accountability systems used in Kentucky and Maryland serve as good models for other states.

References

American Federation of Teachers. (1996). *Making standards matter 1996: An annual fifty-state report on efforts to raise academic standards*. Washington, DC: Author.

Bond, L. & Roeber, E. (1995). *The status of state student assessment programs in the United States*. Washington, DC: Council of Chief State School Officers and North Central Regional Educational Laboratory.

Erickson, R. N., Thurlow, M. L., & Thor, K. (1995). *State special education outcomes 1994*. Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Erickson, R. N., Thurlow, M. L., Seyfarth, A. & Thor, K. (1996). *State special education outcomes 1995*. Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Gandal, M. (1995). *Making standards matter: A fifty-state progress report on efforts to raise academic standards*. Washington, DC: American Federation of Teachers, Educational Issues Department.

Kentucky Department of Education. (No Date). *Kentucky's Learning Goals and Learner Outcomes*.

Kentucky Systems Change Project for Students with Severe Disabilities. (1992). *KIRIS Kentucky Alternate Portfolio Project–Teacher's Guide*. Frankfort, KY: Kentucky Department of Education.

Maryland School Performance Program. (1990). *Learning Outcomes in Mathematics, Reading, Writing/Language Usage, Social Studies, and Science for Maryland School Performance Assessment Program*. Baltimore, MD: Maryland State Department of Education.

Mullis, I. (1990). *The NAEP Guide: A Description of the content and methods of the 1990-1992 assessment*. Washington, DC: National Center for Education Statistics.

Thurlow, M. L., Scott, D. L., & Ysseldyke, J. E. (1995a). *Compilation of states' guidelines for accommodations in assessments for students with disabilities*. Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Thurlow, M. L., Scott, D. L., & Ysseldyke, J. E. (1995b). *Compilation of states' guidelines for including students with disabilities in assessments*. Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Zlatos, B. (1994). Don't ask, don't tell. *The American School Board Journal, 11*, 24-28.

APPENDICES

APPENDIX A

Samples of Released Items and Scoring Guides from
Maryland School Performance Assessment Program (MSPAP)

MSPAP
PUBLIC RELEASE TASK

Salinity

Grade 5

Science
Language Usage

NOTICE

Use of this and all non-secure public release tasks to train educators and to familiarize students with the Maryland School Performance Assessment Program is encouraged. However, copyrighted text material cited on the acknowledgments page is protected by federal copyright law. Before duplication and dissemination of copyrighted material on an extensive basis, consideration should be given to securing appropriate copyright permission.

Maryland State Department of Education
July 1994

State of Maryland
William Donald Schaefer, *Governor*

Maryland State Department of Education
Division of Planning, Results, and Information Management

STATE BOARD OF EDUCATION

Robert C. Embry, Jr., *President*	Baltimore	1995
Edward Andrews, *Vice President*	Rockville	1996
Christopher T. Cross	Chevy Chase	1997
George W. Fisher, Sr.	Denton	1997
Christopher E. Grant	Phoenix	1996
Marvin E. Jones	Glen Burnie	1995
Elmer B. Kaelin	Hagerstown	1995
Rose LaPlaca	Mitchellville	1998
Joan C. Maynard	Linthicum Heights	1996
Harry D. Shapiro	Baltimore	1997
Edmonia T. Yates	Baltimore	1998
Joseph Edmonds, *Student Member*	Baltimore	1995

Nancy S. Grasmick, *Secretary-Treasurer of the Board, State Superintendent of Schools*

A. Skipp Sanders, *Deputy State Superintendent of Schools*

Mark Moody, *Assistant State Superintendent, Division of Planning, Results, and Information Management*

Steve Ferrara, *Director of Maryland School Performance Assessment Program*

Susan Carole Ciotta, *Graphics Coordinator*

Gail Goldberg, *Specialist in Test Development*

The Maryland State Department of Education does not discriminate on the basis of race, color, sex, age, national origin, religion, or disability in matters affecting employment or in providing access to programs. For inquiries related to Departmental policy, contact the Equity Assurance and Compliance Branch.

ACKNOWLEDGMENTS

The Maryland State Department of Education would like to thank:

- The hundreds of local school system staff who are members of the Content Coordinating Teams, the Test Administration Team, the Scoring Advisory Team, and the Task Development Teams that developed these and all MSPAP tasks.
- Its specialists from the Division of Instruction, the Division of Planning, Results, and Information Management, and other divisions for their assistance in the creation of MSPAP.

INTRODUCTION

Background and Purpose of This Document

Maryland public schools have embarked on an important mission: to "re-form" and improve Maryland public schools so that all children can learn, attend schools in which they can progress and learn, and have a real opportunity to learn equally rigorous content. The Maryland State Department of Education launched the Maryland School Performance Program, its strategy for improving public education, in 1989. One component of the Maryland School Performance Program is the state's performance-based assessments, often referred to as the "CRTs" (for criterion-referenced tests). These assessments require students to apply what they know and can do to solve problems, reason, explain, recommend, and display other "higher order" thinking skills. These assessments are officially called the Maryland School Performance Assessment Program (MSPAP).

The primary focus of MSPAP is *school performance*. However, individual student scores from MSPAP are also available. MSPAP assessment tasks assess student performance in grades 3, 5, and 8 in relation to the Maryland Learning Outcomes. These outcomes focus on what students should know and be able to do in reading, writing, language usage, mathematics, science, and social studies. MSPAP tasks and the learning outcomes they assess are sometimes confused with "outcome based education," an approach to teaching, learning, and managing schools which has its share of supporters and detractors. The purpose of this document is to help parents, teachers, students, and other Maryland citizens understand what MSPAP tasks are like.

This Document

This document contains information and materials related to one of nine operational MSPAP tasks that have been selected for public release. These nine tasks were selected to illustrate the types of activities, questions, and responses that MSPAP requires of students.

Contained in this document are actual operational test materials, including:

- *Student Response Book/Answer Book:* Contains questions and other directions to students and space for students to enter their responses.
- *Student Resource Materials Book/Resource Book:* Contains background reading and other information. Only some tasks require such background material.
- *Manipulatives:* Additional materials necessary for tasks (e.g., spinners for the mathematics task "School Fair").
- *Examiner's Manual:* Contains directions to teachers who administer MSPAP, including the directions they read to students verbatim.

Information on scoring these tasks is available in the *Scoring Guide — Introduction, Scoring Tools, and Sample Responses*. This document explains how student responses to the MSPAP are scored. It also contains criteria used to score student responses and sample student responses to all assessment activities in the task.

As you examine this document you will quickly see the complexity of the materials related to each MSPAP assessment task. Because of this complexity, MSDE distributes these tasks at the request of citizens only in conjunction with a brief guided presentation of the materials by an MSDE or local school system educator.

We hope you find the materials interesting and informative. We also expect that you will recognize the power that assessments like MSPAP have for guiding and goading improvements in school performance and student learning and for raising standards for performance for all Maryland public schools.

ANSWER BOOK

Thursday, Task 2
Title: Salinity

INTRODUCTION

Earth is a unique planet of the solar system in that it supports conditions that are necessary for life. One of the basic conditions necessary for life to occur is water. Over 70 percent of our planet is covered with fresh or salt water. Different organisms are found in each of these environments. The Chesapeake Bay is an environment where fresh water from rivers and streams combines with the salty Atlantic Ocean water. This mixture of salt and fresh water creates a solution known as "brackish water."

Today you will investigate how a hydrometer can be used to measure the different levels of saltiness (salinity) in water samples. You will also decide how the hydrometer could be used to establish the correct salinity for an aquarium.

You will have 42 minutes to complete Activities 1 through 7. Refer to the task summary table on the chalkboard; it tells you how much time is suggested for working as a group or as an individual.

1 Building the hydrometer

Groups of four students

Your group will be constructing a hydrometer. A hydrometer is a device that can be used to measure the differences in saltiness between several water samples. To make this tool, your group will need to collect the following materials from the Materials Center:

- one 4-inch piece of a clear drinking straw
- a small amount of clay
- two BBs

Follow these directions to build a hydrometer for your group to use in the next activity:

1. Cut the straw in half so that you have two pieces of equal length. Set one half of the straw aside. It will not be used.
2. Flatten the clay so that it is about 1" x 1/4".
3. Push the cut straw into the clay so that one end of the straw is sealed. Set the excess clay aside.
4. Carefully drop the two BBs into the open end of the straw. Your group has constructed a home-made hydrometer. Your hydrometer should look like the one on the left. You will use it in the next activity.

2 Investigating salinity

Groups of four students

Your group will use the hydrometer to test the saltiness of different samples of water. Your teacher has prepared water samples for you to test. Be sure to label the samples immediately so that you know which is fresh water and which is salt water. Gather the following materials from the Materials Center:

- 1 cup of fresh water
- 1 cup of salty water
- 3 empty cups
- paper towels (just in case . . .)
- 1 ruler
- 1 thin-line permanent marker
- masking tape

Read and follow the directions very carefully.

Step A Label the two water samples. Place the hydrometer (clay end down) in the fresh-water sample and release it. Observe what occurs. If the hydrometer should leak, remove it from the water and re-seal it with the clay. Draw and label your observations in a picture that shows the hydrometer floating in the fresh-water sample.

Step B Place the hydrometer in the salt-water sample and observe what occurs. Draw and label your observations in a picture that shows the hydrometer floating in the salt-water sample.

Step C In your group, decide on a way to quantitatively measure the level at which the hydrometer floats in fresh water. Write a description of your group's method of measurement.

Step D Using the method of measurement your group decided on, measure and record the level at which the hydrometer floats in the salt-water sample. Record your results below.

Now measure and record the level at which the hydrometer floats in the fresh-water sample. Record your results below.

Step E Describe any differences between the two water samples that you observed.

Step F What are some reasons that might explain any observed differences between the salt and fresh water?

3 In a moment you will be mixing equal amounts of fresh and salt water together. Before you do this, think about the hydrometer and the way that it floated in the previous two samples. Predict how the hydrometer might float in the mixture of fresh and salt water. Write your prediction below. Be sure to give a reason for your prediction.

4 Groups of four students

Step A Working with your group, pour one half of your fresh-water sample into an empty cup and an equal amount of your salt-water sample into a second empty cup. Mix the two water samples together in a third cup and test the mixture with your hydrometer. Record your observations and measurements below.

_____

_____

_____

Step B Did your investigation cause you to accept or reject the prediction you made? Explain why, using evidence from your investigation.

_____

_____

_____

5 You have just completed an investigation that involved water with different salinity values. In the next activity you will use this information to solve some problems that might occur when you are keeping animals and plants in an aquarium.

In the Chesapeake Bay, salinity determines the types of animals and plants that can survive in a particular zone. Some types of fish can only be found in areas that have a certain amount of salt in the water. Salinity can be measured in parts per thousand, or "ppt." Higher ppt measurements indicate greater salinity.

Step A

The chart below represents several species of organisms that are common to the bay. It also includes the range of salinity in which the organisms can live. Open your Resource Book to page 10 and use the map of the Chesapeake Bay and the chart below to complete the last column in the chart.

SALINITY SURVIVAL ZONES

Organism	Salinity Range	Zones Where the Organism Can Be Found
Blue Crab	0–30 ppt	
Black Sea Bass	15–30 ppt	
Sea Nettle	7–30 ppt	
White Crappie	0 ppt	
Striped Bass	0–30 ppt	
Common Sea Star	18–30 ppt	
Marsh Periwinkle	0–15 ppt	
Waterweed	0–9 ppt	
Yellow Pond Lily	0 ppt	

Step B

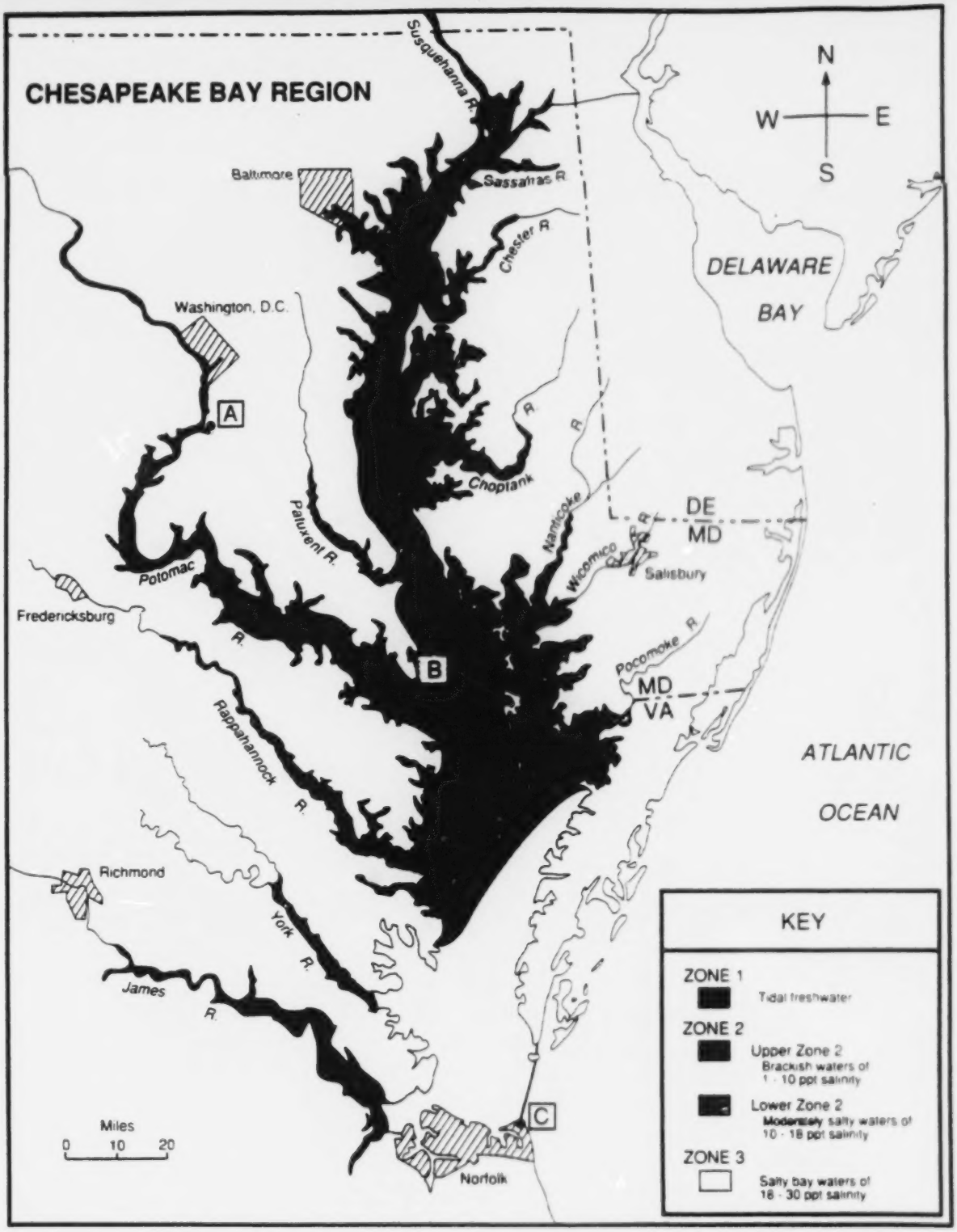
The salt-water aquarium in your school has a salinity range of 16 to 30 ppt. From the list of organisms above, identify the plants or animals that would NOT be able to survive in the aquarium and explain your reasons for not including these organisms. Write your answer below.

6 On a recent field trip to the Chesapeake Bay, your class caught several small black sea bass for the school aquarium. Write a paragraph for your teacher describing how you could use the hydrometer to make sure that these fish stay alive. Use observations and data from what you did today to help you write your response below.

7 Draw a circle around the number below that shows how easy or how hard it was for you to complete the activities in this task.

1	2	3	4	5
Very easy	Somewhat easy	About average	Somewhat hard	Very hard

RESOURCE BOOK



STOP.
Turn to page 37 in your Answer Book.

BEST COPY AVAILABLE

EXAMINER'S MANUAL

Thursday

Secure information pertaining to other tasks

has been removed from this area.

Task 2 Title: Salinity

Total testing time for this task, beginning with oral directions, is **45 minutes.**

Copy the following task summary table onto the chalkboard.

Activity	How Students Work	Suggested Time (Minutes)
1	Group of 4	3
2	Group of 4	15
3	Individual	5
4	Group of 4	5
5	Individual	7
6	Individual	6
7	Individual	1

BEST COPY AVAILABLE

Thursday

MATERIALS/EQUIPMENT NEEDED, LISTED BY ACTIVITY

Activity 1
For each group of four students:

- one 4-inch section of a clear drinking straw
- a small amount of modeling clay, about the size of a regular marble
- two copper BBs, plus extras if students drop theirs

Activities 2 and 4
For each group of four students:

- three 12-oz clear plastic cups (empty)
- one metric ruler
- one thin-line permanent ink marker
- masking tape
- paper towels
- one 12-oz clear plastic cup containing 10 oz of salt water
- one 12-oz clear plastic cup containing 10 oz of fresh water

TEACHER PREPARATION

Prepare a plan for grouping the students in groups of four.

Before administering the task, prepare 1 gallon of salt water according to the following recipe:

- 4 cups of kosher or canning salt
- $\frac{3}{4}$ gallon of water from the tap

Mix salt and water in an empty one-gallon container. Close the top. Shake vigorously. If there is salt left on the bottom, you have made it correctly.

Setting up for the test:

Shake the container of water. Pour approximately 10 oz of the salt water into the clear plastic 12-oz cups. Supply enough water for each group of four students. Do the same with the fresh water. Have supplies set in a Materials Center for the students to get during test administration. Place these in an area labeled "Salt Water Cups" or "Fresh Water Cups." Students will take cups from this area.

SAY **Open your Answer Book to page 32 and follow along as I read aloud.**

Allow time for the students to locate the correct page.

SAY **INTRODUCTION**

Earth is a unique planet of the solar system in that it supports conditions that are necessary for life. One of the basic conditions necessary for life to occur is water. Over 70 percent of our planet is covered with fresh or salt water. Different organisms are found in each of these environments. The Chesapeake Bay is an environment where fresh water from rivers and streams combines with the salty Atlantic Ocean water. This mixture of salt and fresh water creates a solution known as "brackish water."

Thursday

SAY **Today you will investigate how a hydrometer can be used to measure the different levels of saltiness (salinity) in water samples. You will also decide how the hydrometer could be used to establish the correct salinity for an aquarium.**

You will have 42 minutes to complete Activities 1 through 7. Refer to the task summary table on the chalkboard; it tells you how much time is suggested for working as a group or as an individual.

Write the remaining testing time on the chalkboard at appropriate intervals.

Secure information pertaining to other tasks

has been removed from this area.

MSPAP

PUBLIC RELEASE TASK

Salinity

SCORING GUIDE

INTRODUCTION, SCORING TOOLS,
AND SAMPLE RESPONSES

Grade 5

Science
Language Usage

Maryland State Department of Education
July 1994

State of Maryland
William Donald Schaefer, *Governor*

Maryland State Department of Education
Division of Planning, Results, and Information Management

STATE BOARD OF EDUCATION

Robert C. Embry, Jr., *President*	Baltimore	1995
Edward Andrews, *Vice President*	Rockville	1996
Christopher T. Cross	Chevy Chase	1997
George W. Fisher, Sr.	Denton	1997
Christopher E. Grant	Phoenix	1996
Marvin E. Jones	Glen Burnie	1995
Elmer B. Kaelin	Hagerstown	1995
Rose LaPlaca	Mitchellville	1998
Joan C. Maynard	Linthicum Heights	1996
Harry D. Shapiro	Baltimore	1997
Edmonia T. Yates	Baltimore	1998
Joseph Edmonds, *Student Member*	Baltimore	1995

Nancy S. Grasmick, *Secretary-Treasurer of the Board, State Superintendent of Schools*

A. Skipp Sanders, *Deputy State Superintendent of Schools*

Mark Moody, *Assistant State Superintendent, Division of Planning, Results, and Information Management*

Steve Ferrara, *Director of Maryland School Performance Assessment Program*

Susan Carole Ciotta, *Graphics Coordinator*

Gail Goldberg, *Specialist in Test Development*

The Maryland State Department of Education does not discriminate on the basis of race, color, sex, age, national origin, religion, or disability in matters affecting employment or in providing access to programs. For inquiries related to Departmental policy, contact the Equity Assurance and Compliance Branch.

ACKNOWLEDGMENTS

The Maryland State Department of Education would like to thank:

- The hundreds of local school system staff who are members of the Content Coordinating Teams, the Test Administration Team, the Scoring Advisory Team, and the Task Development Teams that developed these and all MSPAP tasks.
- Its specialists from the Division of Instruction, the Division of Planning, Results, and Information Management, and other divisions for their assistance in the creation of MSPAP.

INTRODUCTION TO SCORING GUIDE

The following collection of scoring tools comes from an operational scoring guide used by the team of Maryland teachers who scored this task from the 1994 MSPAP. It is important to remember that such scoring guides are only a supporting resource used to train teachers to score MSPAP. The guides, along with several training sets, are reviewed over two days of training. During this time, the team coordinator and team leader identify essential criteria for each score decision, and carefully explain why the particular sample responses were selected to represent a given score point. Many of the samples in these training materials are not immediately obvious examples of a given score point. They have been selected to permit discussion of a wide range of scoring issues and factors that contribute to a score decision. During training, teachers have an opportunity to compare many sample papers and seek guidance on the score decision process. Only after this training do they score several qualifying sets, packets with pre-established "true" scores against which their performance is compared. While the primary purpose of these qualifying sets is to ensure that teachers can score with sufficiently high exact agreement before beginning the actual scoring process, these responses also provide further opportunities to explore and clarify factors contributing to score decisions.

It is likely, therefore, that some of the scores for the sample responses that follow will not be immediately clear to readers. They all have been selected, however, based on consensus by a team of Maryland educators, backed up by the scoring contractor's senior staff. It will be most helpful if you use them to get a general "feel" for what characterizes both satisfactory and excellent performance. Remember that a score of "1" always means "at least satisfactory." While there may be some ideas or information of merit in some sample "0's," in the judgment of Maryland teachers those responses were not yet satisfactory. They might be too vague, contain too much erroneous information, or require the scorer to make too many inferences about intended meaning. While scoring tools vary in terms of the number of score points that may be assigned, the highest score point is reserved for excellent responses. These are responses that fully address the demands of the particular activity and reflect the conventions of the discipline being assessed.

You will note that some activities are scored more than once, each time with a different scoring tool. This is called "sequential" scoring. This is done when a response may demonstrate a different degree of competency in one outcome area than it might in another which is also being measured by that same activity. Sequential scoring is used in this task for one activity intended to yield both a content area outcome score and a language in use score.

The following is a summary of the assessment activities (i.e., items) in this task and the areas of knowledge and skills assessed by each activity. Scoring tools and actual student responses follow.

Activity	Outcome Assessed
1	DO NOT SCORE*
2A	Processes of Science (S outcome 5)
2B	Processes of Science (S outcome 5)

Activity	Outcome Assessed
2C	Processes of Science (S outcome 5)
	Applications of Science (S outcome 6)
2D	Processes of Science (S outcome 5)
2E	Nature of Science (S outcome 2)
2F	Nature of Science (S outcome 2)
3	Habits of Mind (S outcome 3)
	Processes of Science (S outcome 5)
4A	Processes of Science (S outcome 5)
4B	Nature of Science (S outcome 2)
5A	Processes of Science (S outcome 5)
5B	Concepts of Science/Earth Science (S outcome 1)
	Applications of Science (S outcome 6)
6	Applications of Science (S outcome 6)
6	Language in Use (LU outcome 1)**
7	Transfer Item (degree of difficulty)***

---

* The designation "Do Not Score" indicates that an activity is needed for the "flow" of the task but is not intended to provide a measure of any Maryland learning outcome.

** This activity was scored with the Language in Use rule (0-2 scale). Training on this scale is not task-specific; in this case, readers trained using examples of LU activities from tasks other than Salinity. A copy of this rule is included as an appendix to this guide. Some examples of LU scores for activities from other tasks are included in *Scoring MSPAP: A Teacher's Guide*.

*** At the end of some MSPAP tasks, students are asked to circle the statement that best matches their feelings about how easy or how hard it was to do that task. These are called "transfer items" since they are not scored. Readers simply transfer the information by bubbling in the score sheet to match the number of the statement circled. Data on perceived degree of difficulty can then be scanned and aggregated.

Activity 1 - DO NOT SCORE

Activity 2A

The response demonstrates the ability to employ the language, instruments, methods and materials of science for communicating information.

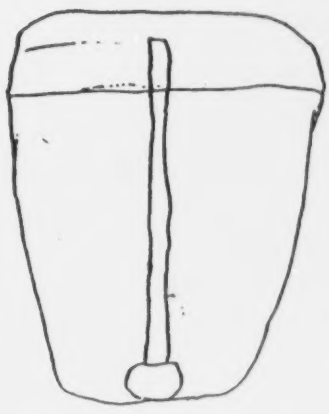
2 = The response includes drawing and written labels: Drawing is labeled. Simply labeling hydrometer as such is a 1, not a 2, because the parts need to be labeled.

1 = The response shows the hydrometer with unlabeled drawing.

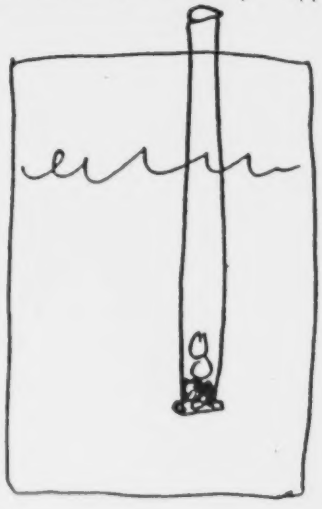
0 = Other

Step A

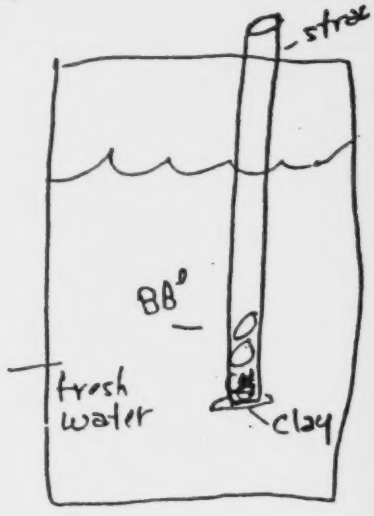
Label the two water samples. Place the hydrometer (clay end down) in the fresh-water sample and release it. Observe what occurs. If the hydrometer should leak, remove it from the water and re-seal it with the clay. Draw and label your observations in a picture that shows the hydrometer floating in the fresh-water sample.



Score = 0



Score = 1



Score = 2

If the picture shows the hydrometer's straw on a slant, this is acceptable because that could be what they observed.

Activity 2B

This response demonstrates the ability to employ the language, instruments, methods and materials of science for communicating information.

2 = The response includes drawing and written labels: Drawing is labeled.

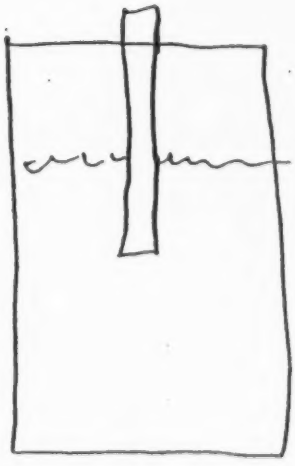
1 = The response shows the hydrometer with unlabeled drawing.

0 = Other

Step B

Place the hydrometer in the salt water sample and observe what occurs. Record your observations in a picture that shows the hydrometer floating in the salt water sample.

Score = 0



Score = 1

Score = 2

Step C

In your group, decide on a way to quantitatively measure the level at which the hydrometer floats in fresh water. Write a description of your group's method of measurement below.

We use centimeters and then we subtract how high the bottom of the straw is from the top

Score = 2

Step C

In your group, decide on a way to quantitatively measure the level at which the hydrometer floats in fresh water. Write a description of your group's method of measurement below.

I am going to put a pencil on it side and see how high it floats

This response is vague (what is the antecedent for it?) but does show a non-standard unit of measure.

Score = 1

Step C

In your group, decide on a way to quantitatively measure the level at which the hydrometer floats in fresh water. Write a description of your group's method of measurement below.

We took the ruler in the cup and measured it.

Score = 0

This response is not clear enough for someone to replicate.

BEST COPY AVAILABLE

Activity 2C

This response demonstrates the ability to apply science and employ the instruments for interpreting information.

2 = The response includes an appropriate means of making the measurement and the use of standard metric units (cm) to measure.

1 = The response includes a description of an appropriate means of measuring but uses a non-standard unit of measuring the level at which the hydrometer floats. (Anything besides cms is considered non-standard: i.e., inches.)

0 = Other

Sample Responses:

2 = Our group decided to measure the distance from the bottom of the floating hydrometer to the table surface. We did this by placing a mark on the glass and then, using the (metric) ruler, we placed the ruler on the table and measured in cm to the top of the mark on the cup. Could measure in inches also.

1 = Our group decided to place a piece of paper next to the glass and made a mark that represents where the hydrometer is floating. (Response should be clear enough that scorer would be able to do it.)

Answer Cues:

Students could decide to measure the amount of the straw that is above the water level and quantify that distance (measurement),

OR

students may choose to measure from the table top (or desk) to the top of the straw and quantify that distance (measurement).

As long as the means of measuring is consistent, it is acceptable.

The answer should deal with the level (height), not slant.

Activity 2D

This response demonstrates the ability to use instruments, methods and materials to measure.

1 = The response includes two measurements recorded in metric or English measure - one for fresh water, one for salt water.

0 = Other

Answer Cues:

The number the students give as a measurement depends on the measuring method.

If the student is consistent and explains method in 2C correctly, then in 2D, if only one measure is labeled, it can still be scored a 1.

Step C

In your group, decide on a way to quantitatively measure the level at which the hydrometer floats in fresh water. Write a description of your group's method of measurement below.

We use centimeters and then we subtract how high the bottom of the straw is from the top.

Score = 2

Step D

Using the method of measurement your group decided on, measure and record the level at which the hydrometer floats in the salt water sample. Record your results below.

12-4½ centimeters = 7½ centimeters

Now measure and record the level at which the hydrometer floats in the fresh water sample. Record your results below.

14-3 centimeters = 11 centimeters

Score = 1

This response gives the results of the method in part C.

BEST COPY AVAILABLE

2 Step D

Using the method of measurement your group decided on, measure and record the level at which the hydrometer floats in the salt water sample. Record your results below.

3 1/2

Now measure and record the level at which the hydrometer floats in the fresh water sample. Record your results below.

2 1/2

Score = 0

If the student is consistent and explains method in 2C correctly, then in 2D only one measure is labeled, it can still be scored a 1.

Activity 2E

This response demonstrates the ability to use scientific knowledge to interpret and communicate information.

2 = The response describes observable differences and includes comparison of recorded measurement. Student should record level.

1 = The response describes observable differences but lacks comparison of recorded measurement.

0 = Other

Sample Responses:

2 = The hydrometer floated higher in the salt water (4.5 cm) than it did in the fresh water (3.0 cm).

1 = The hydrometer floated higher in the salt water than in the fresh water.

1 = We did not observe any difference. (The answer to 2D shows this.)

Activity 2F

This response indicates an ability to provide evidence to support ideas.

1 = The response demonstrates understanding of the connection between saltiness and the floating level of the hydrometer.

0 = The response is a guess, not based on observation.

Sample Responses:

1 = The salt in the water pushed the hydrometer up. Salt made the water "thicker" and the hydrometer couldn't sink into it like it did in the fresh water.

0 = There are no reasons because we did not observe any differences.

Step E

Describe any differences that you observed.

The straw slanted in the salt water and went lower in fresh water

Score = 1

Step F

What are some reasons that might explain any observed differences between the salt and fresh water?

The salt held the straw up

Score = 1

Step E

Describe any differences that you observed.

salt water hydrometer floats more because of salt

fresh water hydrometer does not float it has no salt

Score = 1

Step F

What are some reasons that might explain any observed differences?

salt caused hydrometer to float

no salt made it sink

Score = 0

This is not a reason.

Activity 3

This response demonstrates the ability to make predictions and give evidence to support answers.

2 = The response is a written prediction, supported by evidence (reason).

1 = The response is a written prediction without evidence, **OR** a prediction with evidence that is not logical.

0 = Other

Sample Response:

2 = I predict that the hydrometer will float somewhere in between the fresh and salt water. The mixture should have half as much salt in it, so the hydrometer should float half as high.

Note to scorers: a prediction about the angle or the level is acceptable: any prediction that relates logically to the observation.

3 In a moment you will be mixing equal amounts of fresh and salt water together. Before you do this, think about the hydrometer and the way that it floated in the previous two samples. Predict how the hydrometer might float in the mixture of fresh and salt water. Write your prediction below. Be sure to give a reason for your prediction.

I predict that the hydrometer will float in the middle because the water is being mixed together.

Score = 2

3 it is not going to float as good but it will float

Score = 1

Is floating in a circle

Score = 0

BEST COPY AVAILABLE

Activity 4A

This response demonstrates the ability to measure using *metric units*.

2 = The response includes written observation and measurement.

1 = The response includes written observation and does not include measurement, **OR** measurement with no observation or an observation which is vague.

0 = Other

Answer Cues:

Measure the distance the hydrometer will float in the fresh/salty water mixture and compare it to the measurements in fresh and salty waters.

Students should not include height/level again.
Inches are not acceptable.

Sample Response:

2 = When we placed the hydrometer in the mix of salt and fresh water, it floated about halfway between the two samples. When we measured, the distance that it floated was 3.5 cm.

1 = When we placed the hydrometer in the water, it floated about halfway between the other two samples.

Activity 4B

This response demonstrates the ability to interpret and explain information generated by the exploration of scientific phenomena.

2 = The response accepts or rejects prediction and provides clear evidence.

1 = The response accepts or rejects prediction with general explanation.

0 = Other

Sample Response:

2 = I accept the prediction, because the experiment shows that the hydrometer's level is directly proportionate to the amount of salt in the water. (Proof is given.)

OR

2 = I do not accept the prediction, because the hydrometer's level isnot directly proportionate to the amount of salt in the water. (Proof is given.)

1 = I accept the prediction (no explanation).

4 Groups of four students

Step A

Working with your group, pour one-half of your fresh water sample into an empty cup and an equal amount of your salt water sample into a second empty cup. Mix the two water samples together and test the mixture with your hydrometer. Record your observations and measurements below.

the hydrometer sank below the salt water measurment and above the fresh water measurment

Score = 1

Step B

Did your investigation cause you to accept or reject the prediction you made? Explain why, using evidence from your investigation.

It caused me to accept because I knew that the fresh water made it sink and the salt water made it rise up more so it was between.

Score = 2

4

Step A

it floated 1/2 way in salt and fresh water - 1/2 between 4.2 and 3.6 at 3.9

Score = 2

Step B

Did your investigation cause you to accept or reject the prediction you made? Explain why, using evidence from your investigation.

except it because I predicted write

Score = 1

Scorer must refer to 2D to see if 4A is a 2.

BEST COPY AVAILABLE

Activity 5A

This response demonstrates the ability to organize and display data.

2 = The response correctly identifies all the zones in which the listed organisms live and must include multiple zones. (Upper and lower must be listed.)

1 = The response correctly identifies at least one correct zone for each organism. (Upper and lower is not necessary.)

0 = Other

Answer Cues:

Identifies organisms on the chart with respect to proper salinity ranges and zones.

Activity 5B

This response demonstrates the ability to interpret data and apply science in making personal decisions to solve a practical problem.

2 = The response contains four correct organisms from the list; explanation identifies ppt values of organisms as too low.

1 = The response contains three or four correct organisms from the list and the interpretation of the ppt is vague or not present.

0 = Other

Answer Cues:

Identifies organisms and desired living conditions.

Sample Response:

White Crappie, Marsh Periwinkle, Yellow Pond Lily, Waterweed. These organisms cannot live in water that is as salty as the aquarium water. Even though the Marsh Periwinkle can live in water as high as 18 ppt, the aquarium water could become too salty for them to live.

5 In the Chesapeake Bay, salinity determines the types of animals and plants that can survive in a particular zone. Some types of fish can only be found in areas that have a certain amount of salt in the water. Salinity can be measured in parts per thousand, or "ppt." Higher ppt measurements indicate greater salinity.

Step A The chart below represents several species of organisms that are common to the bay. It also includes the range of salinity in which the organisms can live. Open your Student Resource Materials Book to page 10 and use the chart and your map of the Chesapeake Bay to complete the last column in the chart.

SALINITY SURVIVAL ZONES

Organism	Salinity Range	Zones Where the Organism Can Be Found
Blue Crab	0-30 ppt	ALL
Black Sea Bass	15-30 ppt	ZONE 2 (lower) Zone 3
Sea Nettle	7-30 ppt	ZONE 2 (upper/lower) Zone 3
White Crappie	0 ppt	ZONE 1
Striped Bass	0-30 ppt	ALL
Common Sea Star	18-30 ppt	ZONE 3
Marsh Periwinkle	0-15 ppt	ZONE 1 ZONE 2 (upper/lower)
Waterweed	0-9 ppt	ZONE 1, ZONE 2 (upper)
Yellow Pond Lilly	0 ppt	ZONE 1

BEST COPY AVAILABLE

5

Step A

The chart below represents several species of organisms that are common to the bay. It also includes the range of salinity in which the organisms can live. Open your Student Resource Materials Book to page 18 and use the chart and your map of the Chesapeake Bay to complete the last column in the chart.

SALINITY SURVIVAL ZONES

Organism	Salinity Range	Zones where the organism can be found
Blue Crab	0-30 ppt	Zone 1-3
Black Sea Bass	15-30 ppt	L. Zone 2-3
Sea Nettle	7-30 ppt	4. Zone 2-3
White Crappie	0 ppt	Zone 1
Striped Bass	0-30 ppt	Zone 1-3
Common Sea Star	18-30 ppt	L. Zone 2-3
Marsh Periwinkle	0-15 ppt	Zone 1
Waterweed	0-9 ppt	Zon 1-2
Yellow Pond Lilly	0 ppt	Zone 1

Score = 1

STEP B

The salt water aquarium in your school has a salinity range of 16-30 ppt. From the list of organisms above, identify the plants or animals that would NOT be able to survive in the aquarium and explain your reasons for not including these organisms. Write your answer below.

White Crappie. Marsh Periwinkle Waterweed. Yellow Pond Lilly the reason these organisms could not survive is because they don't have a range of 16-30

Score = 2

BEST COPY AVAILABLE

5

Step A

The chart below represents several species of organisms that are common to the bay. It also includes the range of salinity in which the organisms can live. Open your Student Resource Materials Book to page 18 and use the chart and your map of the Chesapeake Bay to complete the last column in the chart.

SALINITY SURVIVAL ZONES

Organism	Salinity Range	Zones where the organism can be found
Blue Crab	0-30 ppt	1,2,2,3
Black Sea Bass	15-30 ppt	Lower 2.
Sea Nettle	7-30 ppt	2,2,3
White Crappie	0 ppt	1
Striped Bass	0-30 ppt	1,2,2,3
Common Sea Star	18-30 ppt	3
Marsh Periwinkle	0-15 ppt	1,2,2
Waterweed	0-9 ppt	1,2
Yellow Pond Lilly	0 ppt	1.

Score = 1

STEP B

The salt water aquarium in your school has a salinity range of 16-30 ppt. From the list of organisms above, identify the plants or animals that would NOT be able to survive in the aquarium and explain your reasons for not including these organisms. Write your answer below.

Black sea bass, blue crab, sea nettle, Striped bass, common sea star, Marsh Periwinkle,

Score = 0

BEST COPY AVAILABLE

Activity 6

This response demonstrates the ability to use scientific knowledge and available technology to solve a practical problem.

2 = The response describes hydrometer use to determine salinity and indicates how it will be used with the aquarium to match with water from catch site.

1 = The response describes how hydrometer might be used at catch site, **OR** how the hydrometer might be used in the aquarium without referring to the catch site.

0 = Other

Sample Response:

2 = I could use the hydrometer by comparing the height at which it floats in the water we brought the fish home in with the aquarium water. I could keep adding sea salt until they were equal in height. Then, I would know the samples had about the same amount of salt.

6 On a recent field trip to the Chesapeake Bay, your class caught several small black sea bass for the school aquarium. Write a paragraph for your teacher describing how you could use the hydrometer to make sure that these fish stay alive. Use observations and data from what you did today to help you write your response below.

First you could make a hydrometer & put it in the water where you found the fish. Next you could mesure each day how much salt is in the aquarium & add more if you need it. Use a hydrometer to check

Score = 2

6 On a recent field trip to the Chesapeake Bay, your class caught several small Black Sea Bass for the school aquarium. Write a paragraph for your teacher describing how you could use the hydrometer to make sure that these fish are kept in water with the proper salinity. Use observations and data from what you did today to help you write your response below.

I would use The Hydrometer TO MeSure the Salt level in the aquaram. SO The fish can live in The aquarm.

Score = 1

Dear Teachers,
I think our class can use the hydrometer to see if the measurments of the water is right for the fish.

Score = 0

This response does not tell what the hydrometer's use is, it merely rephrases the question. It is not clear whether the student does or does not understand the concept.

BEST COPY AVAILABLE

APPENDIX A

LANGUAGE IN USE RULE

LANGUAGE IN USE RULE

2 points

- ❑ Consistently uses word and sentence order and language choices to express meaning with style and tone. Text conveys uniform impression of correctness* and any errors that are present represent risk-taking.

1 point

- ❑ Sometimes uses word and sentence order and language choices to express meaning with style and tone. Text generally conveys impression of correctness* and errors may or may not represent risk-taking.

0 points

- ❑ Rarely or never uses word and sentence order and language choices to express meaning with style and tone. Text appears error-ridden.

* correct usage, punctuation, spelling, and capitalization.

APPENDIX B

Samples of Released Items and Scoring Guides from
Kentucky Instructional Results Information System (KIRIS)

1994-95 KIRIS OPEN-RESPONSE ITEM
SCORING WORKSHEET

Grade 4 – Reading Question 3

(Academic expectations covered by this item include: 1.2, make meaning.)

3. "Gloria Who Might Be My Best Friend"

   What made Gloria a good friend? Please give examples from the story to explain your answer.

OPEN-RESPONSE 3

Gloria Who Might Be My Best Friend

If you have a girl for a friend, people find out and tease you. That's why I didn't want a girl for a friend — not until this summer, when I met Gloria.

It happened one afternoon when I was walking down the street by myself. My mother was visiting a friend of hers, and Huey was visiting a friend of his. Huey's friend is five, and so I think he is too young to play with. And there aren't any kids just my age. I was walking down the street feeling lonely.

A block from our house I saw a moving van in front of a brown house, and men were carrying in chairs and tables and bookcases and boxes full of I don't know what. I watched for a while, and suddenly I heard a voice right behind me.

"Who are you?"

I turned around and there was a girl in a yellow dress. She looked the same age as me. She had curly hair that was braided into two pigtails with red ribbons at the ends.

"I'm Julian," I said. "Who are you?"

"I'm Gloria," she said. "I come from Newport. Do you know where Newport is?"

I wasn't sure, but I didn't tell Gloria. "It's a town on the ocean," I said.

"Right," Gloria said. "Can you turn a cartwheel?"

She turned sideways herself and did two cartwheels on the grass.

I had never tried a cartwheel before, but I tried to copy Gloria. My hands went down in the grass, my feet went up in the air, and — I fell over.

I looked at Gloria to see if she was laughing at me. If she was laughing at me, I was going to go home and forget about her.

But she just looked at me very seriously and said, "It takes practice," and then I liked her.

"I know where there's a bird's nest in your yard," I said.

"Really?" Gloria said. "There weren't any trees in the yard, or any birds, where I lived before."

I showed her where a robin lives and has eggs. Gloria stood up on a branch and looked in. The eggs were small and pale blue. The mother robin squawked at us, and she and the father robin flew around our heads.

"They want us to go away," Gloria said. She got down from the branch, and we went around to the front of the house and watched the moving men carry two rugs and a mirror inside.

"Would you like to come over to my house?" I said.

"All right," Gloria said, "if it is all right with my mother." She ran in the house and asked.

It was all right, so Gloria and I went to my house, and I showed her my room and my games and my rock collection, and then I made strawberry Kool-Aid and we sat at the kitchen table and drank it.

"You have a red mustache on your mouth," Gloria said.

"You have a red mustache on your mouth, too," I said.

Gloria giggled, and we licked off the mustaches with our tongues.

"I wish you'd live here a long time," I told Gloria.

Gloria said, "I wish I would too."

"I know the best way to make wishes," Gloria said.

"What's that?" I asked.

"First you make a kite. Do you know how to make one?"

"Yes," I said, "I know how." I know how to make good kites because my father taught me. We make them out of two crossed sticks and folded newspaper.

"All right," Gloria said, "that's the first part of making wishes that come true. So let's make a kite."

We went out into the garage and spread out sticks and newspaper and made a kite. I fastened on the kite string and went to the closet and got rags for the tail.

"Do you have some paper and two pencils?" Gloria asked. "Because now we make the wishes."

I didn't know what she was planning, but I went in the house and got pencils and paper.

"All right," Gloria said. "Every wish you want to have come true you write on a long thin piece of paper. You don't tell me your wishes, and I don't tell you mine. If you tell, your wishes don't come true. Also, if you look at the other person's wishes, your wishes don't come true."

Gloria sat down on the garage floor and started writing her wishes. I wanted to see what they were — but I went to the other side of the garage and wrote my own wishes instead. I wrote:

1. *I wish I could see the catalog cats.*
2. *I wish the fig tree would be the tallest in town.*
3. *I wish I'd be a great soccer player.*
4. *I wish I could ride in an airplane.*
5. *I wish Gloria would stay here and be my best friend.*

I folded my five wishes in my fist and went over to Gloria.

"How many wishes did you make?" Gloria asked.

"Five," I said. "How many did you make?"

"Two," Gloria said.

I wondered what they were.

"Now we put the wishes on the tail of the kite," Gloria said. "Every time we tie one piece of rag on the tail, we fasten a wish in the knot. You can put yours in first."

I fastened mine in, and then Gloria fastened in hers, and we carried the kite into the yard.

"You hold the tail," I told Gloria, "and I'll pull."

We ran through the back yard with the kite, passed the garden and the fig tree, and went into the open field beyond our yard.

The kite started to rise. The tail jerked heavily like a long white snake. In a minute the kite passed the roof of my house and was climbing toward the sun.

We stood in the open field, looking up at it. I was wishing I would get my wishes.

"I know it's going to work!" Gloria said.

"How do you know?"

"When we take the kite down," Gloria told me, "there shouldn't be one wish in the tail. When the wind takes all your wishes, that's when you know it's going to work."

The kite stayed up for a long time. We both held the string. The kite looked like a tiny black spot in the sun, and my neck got stiff from looking at it.

"Shall we pull it in?" I asked.

"All right," Gloria said.

We drew the string in more and more until, like a tired bird, the kite fell at our feet.

We looked at the tail. All our wishes were gone. Probably they were still flying higher and higher in the wind.

Maybe I would see the catalog cats and get to be a good soccer player and have a ride in an airplane and have the tallest fig tree in town. And Gloria would be my best friend.

"Gloria," I said, "did you wish we would be friends?"

"You're not supposed to ask me that!" Gloria said.

"I'm sorry," I answered. But inside I was smiling. I guessed one thing Gloria wished for. I was pretty sure we would be friends.

From *The Stories Julian Tells* by Ann Cameron. Copyright © 1981 by Ann Cameron. Reprinted by permission of Pantheon Books, a division of Random House, Inc.

BEST COPY AVAILABLE

Scoring Guide

Score	Description
4	Student gives a relevant response that shows an in-depth understanding that Gloria became a good friend because she respected Julian and didn't make fun of him when he tried to turn a cartwheel and she shared other experiences with Julian in a caring way. Answer gives additional significant examples from the story and/or offers generalizations about Julian's and Gloria's friendship.
3	Student gives a clear response that shows a complete understanding that Gloria became a good friend when she didn't make fun of Julian. Answer gives additional less significant examples from the story.
2	Student gives a weak or limited response that shows a basic understanding that Gloria was a good friend.
1	Student gives a vague or general response, e.g., "she was friendly".
0	Response is totally incorrect or irrelevant.
Blank	Blank/no response

Gloria was a good friend because she

- respected Julian: didn't make fun of him when he tried to turn a cartwheel
- shared experiences: looked at the mother robin, put wishes on the tail of the kite
- had a sense of humor: licked off the Kool-aid mustache

Note: Future open-response questions and scoring guides will quantify what is required of students.

KIRIS Assessment Annotated Response
Grade 4 Reading

Sample 4-Point Response

One thing that made Gloria a good friend was she didn't make fun of Julian when he tried a cartwheel. She just said it takes practice. Another thing that made her a good friend was she told Julian how to wish on something and make it come true. She told Julian to make a kite and wish on something and it would come true. She also wished for her and Julian to become friends. At least, that was what Julian said.

The student shows an understanding that Gloria was a good friend and respected Julian by not making fun of him.

The student realizes that respect is very important if one is to be a good friend.

The student gives other significant examples of friendship, such as sharing and the desire to become friends.

The student predicts that Julian and Gloria will become friends based on the given literary selection.

Sample 3-Point Response

When Julian tried to do a cartwheel, he fell over and Gloria did not laugh. Another reason why she's a good friend is when she told Julian it takes practice.

The student demonstrates Gloria becoming Julian's friend by not making fun of him.

The student might have provided more details in relating the sharing of Kool-Aid, making wishes with the kite, and expressing the desire to become friends.

The student states Gloria indicates she is a friend because she offered Julian advice, which is a less significant example from the story than others that may have been given.

BEST COPY AVAILABLE

KIRIS Assessment Annotated Response
Grade 4 Reading

Sample 2-Point Response

The student shows understanding of the story by stating that Gloria was a friend.

What made Gloria a good friend is she made a friend and she was nice ~~and she had someone to play with~~. What made a good friend is she played with Julian and they flew kites together and drank strawberry Kool-Aid and they drank it.

The student relates less significant events from the story to reflect Gloria and Julian's friendship.

Sample 1-Point Response

The student gives a general answer about friendship, using a minor example from the story.

She likes friends. She plays with them. She plays with her kite and lets them play with it. It goes up in the sky, way high in the sky.

INSTRUCTIONAL STRATEGIES

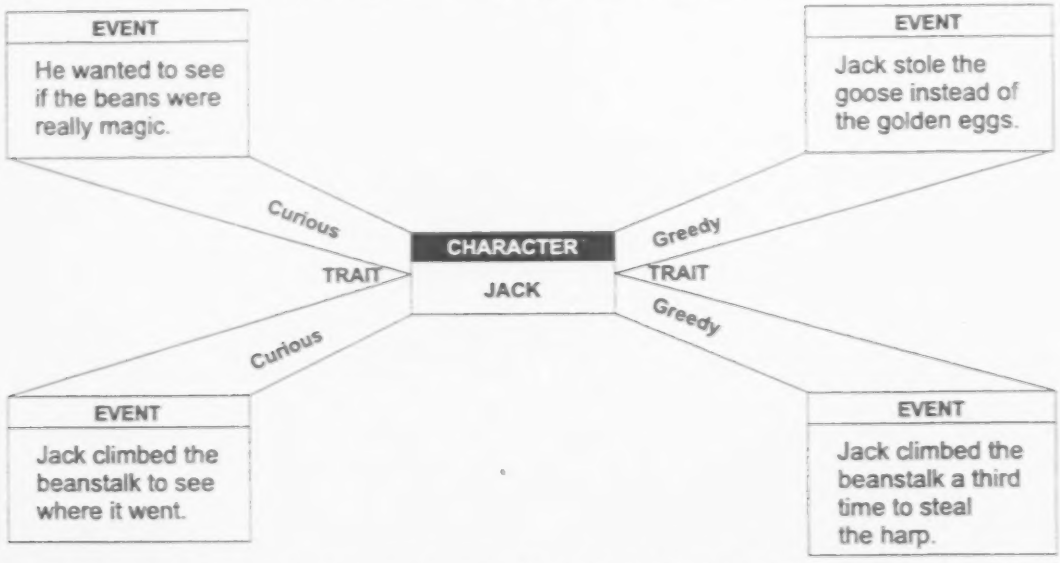
Instructional strategies to assist students in making meaning from literary selections might include the following reading techniques:

1. *Think-Pair-Share (where a question is posed about a reading selection to the whole class, students are asked to think about possible answers individually, then pair with one other student to discuss/refine answer(s), and ultimately share rehearsed answers with the whole class);
2. *Higher-Order Questioning (teacher asks a variety of questions—those with answers stated right in the text, those that require pulling information from various locations in the text, and those with answers that are not in the text and so require students to use their own experiences to answer);
3. *Vocabulary Motor Imaging (vocabulary words are introduced, pantomimes are "imagined" to represent the words, everyone practices the word with the definition and pantomime, and finally text is read containing the new words);
4. *use of a graphic organizer, such as the character trait chart (see below); and
5. provide practice in verbalizing generalizations based on events in a selection. Multiple opportunities should be given to answer questions on a selection and to provide examples to explain/support the response required.

Character Trait Chart

The name of the character from the literary selection is placed in the center of the chart. Adjectives describing the character radiate out from the center box. In each box labeled "event" something from the story is recorded to support the adjective leading to the box.

The following example is generic, based on the familiar story, "Jack and the Beanstalk."



The same type of application could be used with literary characters like Gloria in "Gloria Who Might Be My Best Friend."

*(See *Framework, Vol. II,* for further discussion on graphic organizers.)

BEST COPY AVAILABLE

1994-95 KIRIS OPEN-RESPONSE ITEM
SCORING WORKSHEET

Grade 4 – Science Question 5

(Academic expectations covered by this item include: 2.1, the nature of scientific activity.)

5. The picture shows a full-grown tomato plant. Imagine that you had a real plant, just like the one in the picture, that you could study in any way you wanted so you could completely describe it.

   a. List the properties of the plant and its parts that you think you could describe by just using your senses.

   b. What characteristics of the plant could you describe only by using tools in addition to your senses?

OPEN-RESPONSE 5

Scoring Guide

Score	Description
4	Student gives a comprehensive list of qualitative and quantitative properties of a real tomato plant (like the one in the picture) that could be determined with the unaided senses (color, texture, smell, number of tomatoes). Lists measurements or qualitative data that could be determined with equipment (weight of tomatoes, amount of water in a leaf, height of plant). May use innovative ideas.
3	Student presents a fairly comprehensive list of directly observable properties, but may omit some that are obvious. Also, lists some things that could be measured but lists are not extensive.
2	Student names either the obvious observable properties or the most obvious measurable properties (length, weight). Student may rely on the picture without the concept of the real plant. Student may name tools not directly useful for collecting data. OR Student gives one clearly observable and one clearly measurable property.
1	Student names a general property or two (e.g. size) but does not provide discussion to associate them with observation or measurement.
0	Response is totally incorrect or irrelevant.
Blank	Blank/no response

Note: Future open-response questions and scoring guides will quantify what is required of students.

KIRIS Assessment Annotated Response
Grade 4 Science

The student is able to describe/verbalize with accuracy the properties of the plant that are naturally observable. He/she relates it to a personal experience with reference to taste and smell.

The student lists properties that can be determined with unaided senses.

The student demonstrates the skills needed for scientific activity: collecting data by observing and measuring, noting similarities and differences of properties, and recording/communicating clearly.

Sample 4-Point Response

A. I have a fully grown tomato plant. It is tall and has bright red and green tomatoes on it. The leaves are soft and fuzzy. The red and green tomatoes are smooth and plump. When I ate one of the red tomatoes, it was tender and juicy. It smelled like grass and flowers.

B. I could also measure it to see how tall it is. I would do it with a ruler. I could weigh the tomatoes on a scale.

The student gives a list of qualitative and quantitative properties of the tomato plant.

The student gives quantitative properties by using knowledge of tools of measurement.

Sample 3-Point Response

A. I am supposed to tell you about the tomato plant in the top picture. I think we would have to check by the looks, sounds, smells, and tastes. I do not like tomatoes and I would think that they are sour.

B. I am supposed to tell what characteristics of the plant by using tools. I think that we can dig it up with a hoe to see how long the roots are and measure the fruits or leaves with a ruler. A tomato plant looks green, leafy, with red, yellow, or green fruits. A tomato plant smells bad to me. The only sound would be if you step on it or the wind blows it and the leaves move.

The student lists superficial qualities and does not discuss them.

The student lists some observable properties with the quantitative ones.

BEST COPY AVAILABLE

Sample 2-Point Response

A.
1. Taste it.
2. See if it is ripe.
3. You can't hear it.
4. Smell it.
5. Feel it.

B. The characteristics of the plant I can also describe are how well it grows, if the tomatoes are ripe, cut the tomato open to see if it has any worms or bugs, see if it is healthy, see if there is any water in the dirt, and check to see if it is sprouting. You could count the tomatoes and leaves and check to see how big the tomatoes are. You also can check and count the bugs on the plant. Check to see how bright the tomatoes are. Count up and see if there are any rotten leaves.

The student names obvious observable properties.

The student does not name tools that could be useful for collecting data.

Sample 1-Point Response

taste	smell	see	hear	feel
squishy	sweet	little	nothing	mushy
mushy		round		smooth
smooth		red		
		green		

The student names general properties but does not discuss either measurement or observation specifically.

BEST COPY AVAILABLE

INSTRUCTIONAL STRATEGIES

Instructional strategies that may assist students in collecting data, understanding the nature of scientific activity, and applying higher-order thinking skills include:

1. use of cooperative learning groups;
2. multiple opportunities to conduct and report on scientific investigations, including lab reports;
3. incorporating data collection and interpretation on a regular basis in classroom routines (construct graphs, tables, and charts whenever possible to summarize information);
4. use of real-world application experiences, such as planting a class garden and maintaining records on activities and observable and measurable changes;
* 5. journal writing when applicable to get regular feedback from students in an informal manner;
6. use of student-created projects as they relate to scientific inquiry and plants;
7. multiple opportunities to measure with various tools (both standard and non-standard measurement); and
8. opportunities to work with real plants through hands-on activities.

* Journals are not designed to be "graded," but teacher response in terms of notes to the student or small groups of students is encouraged. Offering students the option of sharing something recorded in their journals could be considered.

Suggested references, in addition to science textbooks and references from the school and local library, are:

a) ***The Budding Botanist,*** an AIMS (Activities Integrating Mathematics and Science) publication, AIMS Education Foundation, Fresno, CA.
b) ***Primarily Plants,*** an AIMS publication, AIMS Education Foundation, Fresno, CA.
c) ***Science and Children*** magazine, an NSTA publication for the elementary level
d) ***Superscience*** magazine, a Scholastic, Inc., publication for the elementary level
e) ***175 Amazing Nature Experiments,*** by Rosie Harlow and Gareth Morgan, Random House, New York, available through NSTA catalog

AIMS books are collections of hands-on activities that can be integrated across the curriculum. All publications contain activities written by teachers and are classroom-tested. Collection and interpretation of data are vital parts of all activities that focus on group cooperation and interaction.

NATIONAL CENTER ON EDUCATIONAL OUTCOMES

University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612/626-1530 • Fax 612/624-0879
http://www.coled.umn.edu/NCEO

The College of Education & Human Development

UNIVERSITY OF MINNESOTA